

THE PSYCHOLOGICAL STUDY OF CONCEPTUAL THINKING¹

By L. S. HEARNshaw

*Department of Psychology, University of Liverpool**I. Current interest in conceptual thinking (pp. 1-3). II. Problems and methods (pp. 3-5). References (p. 5).*

I. CURRENT INTEREST IN CONCEPTUAL THINKING

The psychology of thinking briefly held a prominent place among the interests of psychologists at the beginning of the century, when a group of German investigators, commonly known as the Würzburg school, attempted to analyse human thinking by means of introspection. Prof. Humphrey of Oxford has very ably summarized these investigations in his recent book, *Thinking* (1951), but in spite of his endeavours at rehabilitation the more usual view is that this early work miscarried, primarily because it was attempting an impossible task in trying to unravel complex, rapid and partly subconscious processes by introspective means. At least there can be no doubt that these investigations petered out, and that between the wars experimental investigations into the thought processes of normal adults were few in number. Interest turned to the more experimentally accessible processes of perception and learning, and although intelligence testing thrived prolifically few psychologists, other than Spearman, showed much concern for the nature of the intellect which they were presumably trying to measure.

Since the second World War there has been a growing recognition that the psychology of thinking is not only an important area of psychology but one that can be fruitfully investigated. Hebb (1949), in his already famous book *The Organization of Behaviour*, writes: 'The central problem with which we must find a way to deal... is the problem of thought; some sort of process that is not fully controlled by environmental stimulation and yet cooperates fully with that stimulation. The failure of psychology to handle thought adequately has been the essential weakness of modern psychological theory.' This focusing of interest upon thought processes has been the result of developments in a number of different fields. I propose briefly to review some of the more important of these developments.

In 1934 the Russian psychologist Vigotsky, on the basis of clinical observations and experiments, put forward the theory that a characteristic feature of schizophrenia was loss of the ability to think in abstract concepts and a regression to a more primitive level which he called 'thinking in complexes'. Vigotsky's work was followed up by Hanfmann & Kasanin (1942) in U.S.A. Using Vigotsky's concept formation test they distinguished two main levels and an intermediate level of performance. At the most primitive level objects are grouped in complexes, and these groupings never acquire generality and never transcend the concrete sphere. 'In forming complexes the subject displays a certain passivity... does not stop to deliberate, to formulate a question or hypothesis.' At the highest level grouping is carried out on the basis of abstract general concepts, and subjects show spontaneity, and a capacity for forming hypotheses.

The first paper in a symposium on conceptual thinking, delivered at the Liverpool meeting of the British Association for the Advancement of Science, September 1953.

Remarkably similar general conclusions were reached by Goldstein & Scheerer (1941) working with brain-injured patients. They distinguished abstract and concrete levels of performance. The clinical psychologist has, therefore, become deeply interested in the psychology of thinking and in methods of diagnosing levels of intellectual performance. Unimpaired thinking in persons of average intelligence is marked among other things by the employment of abstract concepts.

Although we must be careful not to identify precisely the levels of regression described by the psychiatrists with development levels, we can, nevertheless, recognize a certain parallelism between them. Piaget, in particular, has made an intensive study of the developing intelligence of the child, and his main conclusions have been conveniently summarized in *The Psychology of Intelligence* (1950), recently translated into English. In his earlier work, *The Language and Thought of the Child* (1926), Piaget was content to recognize two broad levels of thinking, the egocentric level of the child and the socialized intelligence of the adult. His subsequent investigations have forced him to recognize that 'there is a very long way to go from preverbal intelligence to operational thought before reflective groupings may be established'. The two points of the progression are now termed by Piaget sensori-motor intelligence and conceptual intelligence respectively. His researches have done much to expose the complex foundations upon which conceptual intelligence rests, and to revive the interest of psychologists in the higher thought processes.

A further and perhaps decisive factor in forcing the central processes of thinking upon the attention of psychologists has been the realization by experimentalists working largely with animals that perception and learning are not wholly peripheral and self-sufficient processes, but that even at the animal level they are complicated by generalization and transfer, in other words by rudimentary thought processes. No sharp dividing line can be drawn between generalization and transfer at the sensori-motor level and conceptual thinking. As Piaget points out, 'intelligence is not a completely new power superimposed all of a sudden on completely prepared previous mechanisms, but is only the expression of these same mechanisms when they go beyond present and immediate contact with the world'. Fields (1932) in demonstrating that, after training, rats could eventually discriminate any kind of triangle in any position from other figures went on to suggest that they had attained a concept of triangularity. Lashley's (1934) equivalent stimuli, Pavlov's (1927) stimulus generalization, Harlow's (1949) learning sets lead away from the assumption of a sensory dominance of behaviour and towards the acceptance of autonomous central processes. An important aspect of these central processes is that they are concerned with sorting, classifying, ordering, abstracting—with the operations, therefore, that are characteristic of conceptual thinking. As Hebb has rightly seen, experimental psychology moves towards and requires a theory of thought.

Nor have movements outside psychology itself been without their influence. Craik (1943) was one of the first psychologists to grasp the implication for psychology of the development of mechanical and electrical devices capable of receiving, sorting and using information, and in his book, *The Nature of Explanation*, he treats the brain as in some ways comparable to such devices. The attempt to reproduce mechanically some of the intellectual processes of organisms has necessitated a much more exacting analysis of these processes, and this has again brought to the fore some of the cardinal features of conceptual thinking. For example, when Grey Walter (1953) constructed an electrical

mechanism capable of simple learning he found that 'the several learning circuits are really extracting from a selection of events the features that are common to them in time and space—they are in effect recognizing a pattern'. Thus even in quite simple processes there are important analogies to conceptual thinking.

Finally, the philosophers have been approaching what are to them age-old problems with renewed vigour and freshness. Perhaps two recently published works deserve special mention. First, Susanne Langer's (1942) *Philosophy in a New Key*, a brilliant study of symbolism and its functions. 'Symbols', according to her, 'are not proxy for their objects, but are vehicles for the conception of objects... so our interest in the mind has shifted more and more from the acquisition of experience, the domain of sense, to the uses of sense-data, the realm of conception and expression'. The second book is Price's (1953) *Thinking and Experience* which is described by its author as 'an examination of the nature of conceptual cognition, its relation to sense experience on the one hand, and to symbols (verbal and non-verbal) on the other'. Price's analysis of conceptual thinking and his dispositional version of conceptualism are of the highest interest to psychologists. It is indeed in general tenor not unlike accounts given by psychologists themselves, for example by Ribot in his *L'évolution des idées générales* in 1897 or by Wolters in 1933.

Thus by a variety of routes, and for a variety of reasons, the contemporary psychologist has reached the position where he probably would not dissent from William James (1890), who described conception as 'the most important of all features of our mental structure'.

II. PROBLEMS AND METHODS

What are the main problems which confront the psychologist in this field? What methods can he use in his attack on these problems? Is anything of value likely to emerge?

But before we proceed to consider these questions, briefly what is a concept? For the logician abstract general concepts are 'signs pointing to invariant relations, i.e. relations which remain identical despite the variations of the material in which they are embodied. Significant concepts enable us to arrange in order and hold together diverse phenomena because of some real unity of process or relation which constitutes an element of identity between them' (Cohen, 1946). For the psychologist a concept is 'a habit, an organized memory' (Ribot, 1897), or perhaps better as Price (1953) puts it, 'a recognitional capacity'. 'A concept is not before the mind', adds Price, 'as an object of inspection. It is at work in the mind... it shows itself not as a detachable item of mental furniture, but rather as a guiding force, determining the direction which the series of presented particulars takes.' Conceptual dispositions may manifest themselves in a great many different ways, and at all levels from the bare capacity to recognize instances in varying contexts to full blown concepts at a high level of abstraction. Taking Humphrey's (1951) definition we may describe conceptual thinking as 'The activity whereby an organism comes to effect a constant modification towards an invariable feature or set of features occurring in a variable context'. But for the psychologist a merely verbal definition is not enough, and, as Heidebreder (1946, 1947-9) has rightly seen, a precise operational definition of conceptual activity which may serve as criterion of its presence is essential.

The psychologist's problems must be clearly distinguished from the traditional problems of the philosopher. The philosopher is mainly concerned with the ontological and logical status of universals. The psychologist is concerned with the *capacity* for conceptual

thinking, i.e. for grasping universals. It is only comparatively recently that the psychologist's problem has crystallized. The earlier psychological inquiries were concerned with a phenomenological analysis which was in practice difficult to dissociate from philosophical questions. The question that Aveling asked in 1912, for instance, was 'What is discoverable in consciousness when we think the universal?', and similarly Fisher in 1916 asked 'What are the events which transpire in an observer's consciousness after he has been confronted with a task whose successful performance necessitates the act of generalizing abstraction?' Hull's investigation in 1920 marks an important milestone in the progress of conceptual research. He concentrated on the functional and quantitative aspects of the problem which he pointed out were largely neglected in the earlier introspective and qualitative studies. He employed a small group of abnormal subjects in a subsidiary experiment to investigate individual differences in concept formation and thus initiated research into concept formation as a capacity.

The job of the psychologist, therefore, is to investigate conceptual thinking as a capacity of the organism. The field can be broken down into a number of specific problems. From what beginnings and through what stages, first, does this capacity evolve both in the individual and in the race? This is clearly the problem with which the genetic studies of Piaget are primarily concerned. He has traced the development of conceptual thinking in the child not only in respect of conceptual thinking in general, but in relation to specific key concepts such as those of number, space, time, causality, chance and so on. Anthropological investigation and the study of primitive language provide comparable though less adequate data about the phylogenetic development of conceptual thinking. For the beginnings of conceptual thinking we must have recourse to animal studies for, as Price (1953) points out, there is reason to hold that in a sense conceptual thinking is 'coterminous with conscious life itself'. The phenomena of generalization, transfer and learning sets are met with in the study of animal behaviour, and with Hayek (1952) we may perhaps regard the operation of classifying as a basic feature of nervous mechanisms. Evidently a primary task for the psychologist is to map in the greatest detail the whole evolutionary course of conceptual thinking, and, if the term is to be used as broadly as Price suggests, to distinguish levels of complexity. This provides among other things both the best safeguard against the unduly simplified accounts and explanations of the arm-chair philosopher, and the necessary material for an analysis of the underlying mechanisms.

Probably no analysis of these mechanisms is going to be satisfactory unless closely tied to neurophysiological facts, and the precise neurophysiological explanation of these complex events still seems some way distant, though as Hebb's speculations suggest the gap is less wide than it was. Phenomenological analysis by introspective means has proved unreliable in practice, and is doubtful in principle. But a behaviouristic analysis of performance under systematically varied conditions such as that carried out by Hull, Heidbreder (1946, 1947-9) and Reed (1946-50) in U.S.A. may indirectly throw light on the mechanisms of conceptual thinking. A number of interesting conclusions have already emerged from this type of research, though at present they have not been integrated into any systematic behaviour theory embracing the whole range of perception, learning and response, nor have they been linked to neurological events. Such an integration is clearly another of the major tasks of the psychologist.

A third group of problems relates to the functional significance of conceptual development. What does this capacity enable its possessor to do? Is it the precondition of

flexibility of behaviour, of delayed response, of the development of symbolic systems and language? What, in other words, is its biological value as a mode of adaptation?

Fourthly, what is the place of the capacity to conceptualize in the structure of human abilities? Is it the same thing as, or an aspect of, the general factor of intelligence? Is it linked in any way to the verbal factor and linguistic ability? Or to the spatial factor and the ability to appreciate patterns? Again, if rightly described by Price as a 'recognitional capacity', is it related somehow to some memory factor? There is little evidence at present enabling us to answer these questions, but they are questions of the kind that should prove amenable to psychological techniques.

Finally, the psychologist must map the course, and if possible elucidate the causes, of the breakdown of the capacity for conceptual thinking and its regression to lower levels in psychotic conditions, in old age, and as a result of brain injuries.

None of these tasks can be satisfactorily carried out without clear-cut operational definitions of conceptual capacities or without adequate tools. If the psychologist is interested in conceptual thinking as a capacity, the essential tool is one that measures this capacity—in other words a test or tests of known reliability and content. In view of the criticisms which have recently been levelled at tests as *ad hoc* instruments of no theoretical value it is perhaps important to insist that in research into human capacities they are the indispensable tools. There may be methodological problems in their use and interpretation (of which tools is this not true?), but they are not incapable of logical justification or scientific refinement. Existing tests of conceptual thinking, however, leave much to be desired. The Vigotsky test and the Goldstein-Scheerer sorting test have never been adequately standardized or analysed, and as Semeonoff & Laird (1952) point out in their recent study of the Vigotsky test it is practically impossible to demonstrate reliability. The main weakness both of Piaget's development studies, and the clinical studies of conceptual deterioration, is their lack of adequate and reliable measures. The first and most important task confronting the psychologist in the field of conceptual thinking is the construction of better tests. This is a prerequisite not only for research, but also for clinical application and diagnosis.

The practical value of investigations into conceptual thinking can hardly be doubted. The functioning of the normal adult mind is characteristically conceptual: our understanding both of the intellectual development of the child, and of the pathological regressions of later life depends upon our insight into the nature of conceptual thinking. Any techniques for assessing conceptual ability would have an immediate practical value for child and clinical psychologists. But over and above this the investigation of conceptual thinking should throw light on the nature of cerebral processes and of human intelligence, and assist in elucidating some of the central theoretical problems of psychology.

REFERENCES

- AVELING, F. (1912). *On the Consciousness of the Universal and the Individual*. London.
 COHEN, M. R. (1946). *A Preface to Logic*. London.
 CRAIK, K. J. W. (1943). *The Nature of Explanation*. Cambridge.
 FIELDS, P. E. (1932). The development of the concept of triangularity by the white rat. *Comp. Psychol. Monogr.* no. 42.
 FISHER, S. C. (1916). The process of generalising abstraction and its product, the general concept. *Psychol. Monogr.* no. 90.
 GOLDSTEIN, K. & SCHEERER, M. (1941). Abstract and concrete behaviour. *Psychol. Monogr.* no. 239.

- HANFMANN, E. & KASANIN, J. (1942). *Conceptual Thinking in Schizophrenia*. New York.
- HARLOW, H. F. (1949). The formation of learning sets. *Psychol. Rev.* LVI.
- HAYEK, F. A. (1952). *The Sensory Order*. London.
- HEBB, D. O. (1949). *The Organisation of Behaviour*. New York.
- HEIDBREDER, E. F. (1946). The attainment of concepts. I-II. *J. Gen. Psychol.* XXXV.
- HEIDBREDER, E. F. (1947-9). The attainment of concepts. III-VIII. *J. Psychol.* XXIV-XXVII.
- HULL, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychol. Monogr.* no. 123.
- HUMPHREY, G. (1951). *Thinking*. London.
- JAMES, WM. (1890). *Principles of Psychology*, vol. I, p. 460.
- LANGER, S. (1942). *Philosophy in a New Key*. Harvard.
- LASHLEY, K. S. (1943). Nervous mechanisms in learning. In Murchison's *Handbook of Experimental Psychology*.
- PAVLOV, I. P. (1927). *Conditioned Reflexes*. Engl. tr.
- PIAGET, J. (1926). *The Language and Thought of the Child*. Engl. tr. London.
- PIAGET, J. (1950). *The Psychology of Intelligence*. Engl. tr. London.
- PRICE, H. H. (1953). *Thinking and Experience*. London.
- REED, H. B. (1946). Factors influencing the learning and retention of concepts. I-V. *J. Exp. Psychol.* XXXVI-XL.
- RIBOT, T. (1897). *L'évolution des idées générales*. Paris.
- SEMEONOFF, B. & LAIRD, A. J. (1952). The Vigotsky test as a measure of intelligence. *Brit. J. Psychol.* XLIII.
- VIGOTSKY, L. (1934). Thought in Schizophrenia. *Arch. Neurol Psychiat., Lond.*, XXXI.
- WALTER, G. (1953). *The Living Brain*, p. 130. London.
- WOLTERS, A. W. (1933). On conceptual thinking. *Brit. J. Psychol.* XXIV.

(Manuscript received 9 October 1953)

CHANGES OF PERFORMANCE WITH AGE ON SIMILAR TASKS OF VARYING COMPLEXITY

By HILARY M. CLAY

*The Nuffield Research Unit into Problems of Ageing, The
Psychological Laboratory, University of Cambridge*

I. *Introduction* (p. 7). II. *Method* (pp. 7-8). III. *Results* (pp. 8-12): (i) *Accuracy* (pp. 8-10); (ii) *Number of alterations* (pp. 10-11); (iii) *Time taken* (pp. 11-12). IV. *Discussion of results* (pp. 12-13). V. *Summary* (p. 13). *References* (p. 13).

I. INTRODUCTION

Comparisons between young and old on tests of mental ability have shown a decline of performance with age. This deterioration was at first attributed solely to the time limits of the tests used. Lorge (1936) stated that 'speed obscures mental power' which 'probably does not deteriorate with age'. This view was supported by Miles (1938) with the reservation that 'older adults maintain their ability as long as they continue mental practice'. Lawton (1943) expressed his opinion more cautiously: 'If we eliminate speed and consider only mental power, the difference between older and younger individuals is reduced.'

Although the slowing of receptor and effector processes associated with age is the main cause of decline of performance on tests in which success depends on speed, an explanation is still to be sought for the differences observed between younger and older subjects on untimed tests (for example, Raven, 1948; Foulds & Raven, 1948). Other factors, such as length of problem and degree of familiarity with test material, have been examined, but unfortunately mainly in relation to speed. The present experiment is an attempt to isolate one factor, that of complexity in terms of the number of items with which a subject has to deal simultaneously but *at his own speed*, and to study the effect of varying complexity on ability to solve a problem task as one gets older.

II. METHOD

Description of task

Subjects were required to arrange counters, each stamped with a number from 1 to 4, in rows and columns of squares on a board so that the sum of the numbers on the counters in each row and each column equalled the marginal total printed for each on the board.

Four problems were used, varying in the number of counters given, with corresponding changes in the number and value of the rows and columns (see Fig. 1).

In each task the display was similar. The boards were made of white cardboard on which horizontal and vertical black lines were drawn 3.5 cm. apart. The white ivory counters were 2.8 cm. in diameter, and on each was engraved in black a number 1.3 cm. high. They were stacked in a wooden rack according to their values.

The instructions were identical for the four problems. The subjects were told that they should place one counter in each square and that the numbers on the counters should add up, both across and down, to the total given for each row and each column of squares.

			10
			9
			8
10	9	8	

(a) Problem 1

Counters:

Value	1	2	3	4
Number	1	1	4	3

				13
				12
				11
				10
13	12	11	10	

(b) Problem 2

Counters:

Value	1	2	3	4
Number	2	3	6	5

					14
					13
					12
					11
					10
14	13	12	11	10	

(c) Problem 3

Counters:

Value	1	2	3	4
Number	7	6	7	5

						15
						14
						13
						12
						11
						10
15	14	13	12	11	10	

(d) Problem 4

Counters:

Value	1	2	3	4
Number	14	10	7	5

Fig. 1. The boards used in the four problems, showing numbers and values of marginal totals and of counters presented.

The subjects

One hundred and twenty-eight subjects were tested. They were divided into two groups according to age:

Younger group. Sixty-four naval ratings, mean age = 20.9 years (range 18–24 years).

Older group. Sixty-four pensioners, mean age = 67.5 years (range 55–78 years).

Both these groups were divided into four equal sub-groups, to each of which one of the four problems was presented.

All the older subjects, who volunteered to do the experiment, maintained a wide range of interests, although retired from active work.

III. RESULTS

(i) Accuracy

The achievements of the older and younger groups were approximately equal on Problem 1. The number of subjects in both groups who solved the problem fell as the number of counters was increased (see Table 1). The fall was more marked for the older groups. The five older subjects who failed to solve Problem 2, nevertheless managed to get all but two of the marginal totals correct, and the differences between the younger and older groups

on this problem and on Problem 3 were not significant. On Problem 4, however, the extent of the errors made by the older group was significantly greater than that of the younger ($P < 0.05$).¹

Table 1. *Distribution of subjects according to number of incorrect marginal totals*

		No. of incorrect totals											
		0	2	3	4	5	6						
Problem 1		0	2	3	4	5	6						
Younger group	15	1	—	—	—	—	—						
Older group	14	2	—	—	—	—	—						
Problem 2		0	2	3	4	5	6	7	8				
Younger group	15	1	—	—	—	—	—	—	—				
Older group	11	5	—	—	—	—	—	—	—				
Problem 3		0	2	3	4	5	6	7	8	9	10		
Younger group	11	2	1	2	—	—	—	—	—	—	—		
Older group	7	5	1	2	1	—	—	—	—	—	—		
Problem 4		0	2	3	4	5	6	7	8	9	10	11	12
Younger group	12	—	1	—	—	2	1	—	—	—	—	—	—
Older group	5	3	—	3	1	2	—	—	1	—	—	—	1

Table 2. *Analysis of errors*

		Total no. of subjects making errors	Type of error			No attempt to correct
			Simple correction needed	Complex correction needed	Failure to place all counters	
Problem 1						
Younger	1	1	—	—	—	
Older	2	2	—	—	—	
Problem 2						
Younger	1	1	—	—	—	
Older	5	2	3	—	—	
Problem 3						
Younger	5	1	4	—	—	
Older	9	2	3	—	4	
Problem 4						
Younger	4	—	3	1	—	
Older	11	1	4	4	2	
Totals						
Younger	11	3	7	1	—	
Older	27	7	10	4	6	

The errors made may be divided into four types:

(a) A simple error involving either two rows or two columns, which is easily corrected by a direct exchange of two counters. This error may be due to carelessness in checking the totals, but there was some evidence from the subjects' comments that it occurred because they did not realize that by altering the positions of two counters within, for instance, a row, they could correct an error in a column without altering the total of the row. As Table 2 shows, the older group made slightly more errors of this type than the younger.

(b) A more complex error involving two or more rows and columns, which requires not only the inter-change of counters in the incorrect rows and columns, but frequently a transference of the error to a correct row or column before a successful solution can be reached. No errors of this type appeared on Problem 1, and on Problem 2 they were made only by older subjects. Subjects in both groups made errors of this type in Problems 3 and 4.

(c) Failure to complete the task. This type of error only occurred in Problem 4. Four of the older subjects compared to one of the younger, failed to place all the counters on

¹ The tests of significance used were those based on the ranking method developed by Kendall (1948).

the board. They had managed to arrange some of the rows and columns correctly, and then realized that with the counters they had left it would be impossible to complete the problem without altering the positions of those already on the board. This they were unwilling to do.

(d) Although this type is not exactly an error, it is a failure to solve the problem by making no effort to correct mistakes. This type of response was made only by older subjects in Problems 3 and 4. These subjects knew they had made errors, but refused to attempt to correct them on the grounds that either it would take too long or it would upset the correct rows and columns.

(ii) *Number of Alterations*

The number of alterations made in the position of each counter *after* it had once been placed on the board was recorded. This measure gives some indication of attempts to correct errors and of flexibility of methods of tackling the task.

The only significant difference between the two groups in the number of alterations made was in Problem 3, where the younger subjects made more alterations than the older ($P < 0.05$) (see Table 3).

Table 3. *Distribution of subjects according to number of alterations made, and the mean number of alterations for each group*

No. of alterations made	Problem 1		Problem 2		Problem 3		Problem 4	
	Young	Old	Young	Old	Young	Old	Young	Old
0	—	2	3	1	—	1	1	—
1-4	3	2	3	4	1	4	—	2
5-16	7	7	5	3	3	3	3	4
17-64	6	5	5	6	9	7	8	8
Over 64	—	—	—	2	3	1	4	2
Mean	12.4	12.6	13.8	25.9	38.2	20.1	37.1	38.7

As the number of alterations was closely linked with the way in which difficulties were overcome during the performance on the task, it is worth while considering the different methods employed:

(a) Perhaps the most drastic way of coping with difficulties was that adopted by six of the sixty-four younger subjects, and by two of the sixty-four older subjects. They removed all the counters from the board and started again. The two older subjects were in the group doing Problem 4, and one of them replaced the counters in almost identical positions the second time. He was not adopting a new method. On the other hand, the younger subjects took the opportunity of trying a new approach to the task.

(b) Another method involved the exchanging of all the counters in one row, for example, with all those in another, even if this meant that counters of the same value occupied identical positions. This was one way of overcoming the difficulty of correcting errors when no direct exchange of two counters was possible. Only three younger and four older subjects adopted this method.

(c) The most usual way of making alterations was by moving single counters at a time. Often subjects took a counter of the required value from another position on the board, regardless of the effect such a move would have on the solution as a whole. This method involved a number of useless alterations; as one subject remarked: 'All that juggling and it's still wrong.'

These methods suggest that some subjects were tackling the problem in a series of trial-and-error attempts. While it is true to say that a final solution was often reached after a number of piecemeal alterations had been made, an analysis of the initial methods adopted showed that, in the main, subjects approached the problem with some plan, however vague.

The older subjects chose methods similar to those of the younger, but were less successful in carrying them out as the number of counters increased. Fifteen of the sixty-four older subjects compared to three of the younger made an error in the arrangement of the first marginal total completed. Seven of these fifteen older subjects were doing Problem 4. On Problems 3 and 4, the older subjects who started with the intention of arranging the counters so that they added up in both rows and columns, soon forgot or neglected one dimension and concentrated on the other, without correcting their errors before they finished. Or, if they had adopted the method of arranging the counters along one dimension before attempting to fit in with the totals in the opposite direction, they would make the alterations necessary to correct for the other set of totals without any regard to the effect these made on the previously correct arrangement. They treated the problem as two separate tasks.

Table 4. *Distribution of subjects according to times taken, and the mean time for each group*

Seconds	Problem 1				Problem 2				Problem 3				Problem 4			
	Younger		Older		Younger		Older		Younger		Older		Younger		Older	
	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I
	1	1	6	—	3	—	—	1	—	—	—	—	—	—	—	—
Under 120	1	1	6	—	3	—	—	1	—	—	—	—	—	—	—	—
121- 240	7	—	2	—	5	1	2	1	2	—	—	2	—	—	—	—
241- 480	6	—	5	1	5	—	4	2	3	2	3	3	2	—	—	2
481- 960	1	—	1	1	2	—	—	1	2	—	4	1	3	3	5	2
961-1920	—	—	—	—	—	—	4	—	4	2	—	3	7	1	—	4
Over 1920	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	3
Mean time	253		241		256		681		887		579		1007		1166	

C=Correct solution. I=Incorrect solution.

The majority of the older subjects carried on a running commentary throughout, and resorted to such *aides-mémoires* as moving a counter slightly out of place or keeping one finger on the board to indicate an incorrect row or column, while they checked the rest. One or two of the younger subjects counted aloud, but only during the final check of marginal totals when they seemed to use verbal cues in much the same way as the older subjects made use of spatial ones.

(iii) Time taken

Although the subjects were told that there was no time limit, the total time spent on the task was recorded (see Table 4).

Some unpublished results by Kay indicate that speed of performance falls more strikingly with age at difficult than at easy tasks. On the other hand, Birren, Allen & Landau (1952) found that, under limited-time conditions, older subjects were slower in adding digits than younger subjects, and were relatively slowest for the very short or easy problems.

The present results, shown in Table 4, do not entirely agree with either of these previous findings. The average times spent on all the problems, except Problem 2, by younger and

older groups were similar. The older subjects took significantly longer than the younger to complete Problem 2 ($P < 0.05$).¹

A possible reason for these results is found when time taken is considered in relation to difficulty and achievement. Most of the younger subjects on all the problems and of the older subjects on Problems 1 and 2 who took a relatively long time produced correct solutions. This is consistent with the well-known principle that as difficulty (due either to rising demands of the task or to falling capacity of the subject) increased, success can be maintained up to a point by the expenditure of more time. With the older subjects on Problems 3 and 4 this principle seems to break down, and most of those who take a long time fail to produce a correct solution. It would seem that age and the demands of the problems together have produced such a degree of difficulty that for many subjects no amount of extra time they are likely to be able to spend will achieve success. As a result they either plod on unsuccessfully for a long time, or quickly realize that the task is hopeless and give up.

IV. DISCUSSION OF RESULTS

The ability of older subjects to solve the problems is clearly related to the number of counters involved.

The older subjects were able to do the easiest problem, with nine counters, as well as the younger.

There was, however, a progressive breakdown of performance on the part of the older subjects, resulting from the increasing complexity:

- (a) Sixteen counters: the older subjects required *more time* than the younger.
- (b) Twenty-five counters: the older subjects were *less active* in correcting for errors.
- (c) Thirty-six counters: the older subjects were *less accurate*, and 25% of the older group *failed to complete* the task.

Each of these differences was significant.

This breakdown cannot have been entirely due to the decline in accuracy with age in adding as shown by Birren & Botwinick (1951) and by Birren *et al.* (1952). The numbers of errors made in Problems 3 and 4 of this experiment are greater than the numbers made by their subjects adding five and six digits respectively. At the same time, it is clear that the errors remaining in the solutions were not mere errors of addition, because those which were made in the early stages of the task were later corrected—those which remained were recognized but were left because the subjects were unable to correct them.

It seems, therefore, that it was the increasing amount of material presented which caused older subjects more and more difficulty. They could readily *perceive* and *understand* what they were required to do. Failure appears to have been due to changes with age in the amount of data that could be organized together. They were unable to build a solution in a logical way with the more complex problems. Confusion was apparent. They could see their errors, but were unable to make the necessary connexions between errors to arrive at a correct solution. The younger subjects encountered similar difficulties in coping with the increased data, but were methodical in dealing with them. They followed their initial plan more closely, whereas the older subjects became muddled and then neglected some parts of the task.

¹ In a control experiment, no difference was found between the times taken by subjects of different age groups to place twenty-five blank counters in squares on a board.

As the complexity of the tasks was increased, the decline of ability with age was revealed in different ways. This serves to emphasize the necessity, which has been found in previous experiments on ageing (Welford *et al.* 1951), of considering several aspects of performance together when comparing different age groups.

It is perhaps tempting to see these results as linked with a failure of short-term memory, which would make it difficult to relate one piece of evidence in a complex pattern with another. This approach has already been suggested by Welford *et al.* (1951, p. 101) as an explanation of a different type of problem-solving experiment. They point out that the failure may not necessarily be of the memory processes themselves. Any general slowing would be likely to increase the load upon short-term memory, in the sense that whatever was carried would have to be carried for a longer period before it could be used.

V. SUMMARY

In this experiment an attempt was made to study the effects of changes in the number of items to be dealt with simultaneously but without time-stress on the ability of older compared to younger subjects to solve a problem task. Sixty-four subjects under 25 years of age and sixty-four subjects over 55 years of age were tested. They were required to arrange counters of different values in rows and columns on a board, in such a way that the sum of the counters equalled a marginal total given for each row and each column. Four problems were used, in which nine, sixteen, twenty-five and thirty-six counters respectively were presented. Each subject was given one problem only. The results showed that the performances of the two groups were similar on the simplest problem but, as the complexity of the task increased, the older subjects had greater difficulty compared with the younger.

I wish to thank the members of the Cadbury Pensioners' Club who acted as subjects, and also Cadbury Brothers Limited who kindly provided facilities for testing. The naval ratings took part in the experiment through the courtesy of the Admiralty.

My sincere thanks are due to Mr A. T. Welford, Director of the Unit, for his valuable criticism throughout.

REFERENCES

- BIRREN, J. & BOTWINICK, J. (1951). Rate of addition as a function of difficulty and age. *Psychometrika*, **16**, 219-32.
- BIRREN, J., ALLEN, W. R. & LANDAU, H. G. (1952). Speed, accuracy and difficulty. *Rep. psychometric lab. Univ. Chicago*, no. 76.
- FOULDS, G. A. & RAVEN, J. C. (1948). Normal changes in the mental abilities of adults as age advances. *J. Ment. Sci.* **xciv**, 197.
- KENDALL, M. G. (1948). *Rank Correlation Methods*. London: Charles Griffin and Co.
- LAWTON, G. (1943). Aging mental abilities and their preservation. In *New Goals for Old Age*, ed. by G. Lawton, pp. 11-33. New York: Columbia University Press.
- LOGE, I. (1936). The influence of the test upon the nature of mental decline as a function of age. *J. Educ. Psychol.* **xxvii**, 100-10.
- MILES, W. R. (1938). Psychological aspects of aging. In *Problems of Aging*, ed. by E. V. Cowdry. Baltimore: Williams and Wilkins.
- RAVEN, J. C. (1948). The comparative assessment of mental ability. *Brit. J. Psychol.* **xxix**, 12.
- WELFORD, A. T. *et al.* (1951). *Skill and Age*. Oxford University Press for the Nuffield Foundation.

(Manuscript received 5 May 1953)

MEMORY MECHANISMS AND THE THEORY OF SCHEMATA¹

BY R. C. OLDFIELD

Department of Psychology, Reading

I. *Introduction* (pp. 14–15). II. *The character of recall* (pp. 15–16). III. *Circuitual storage devices* (pp. 16–18). IV. *Memory systems based upon circuitual storage elements* (pp. 18–19). V. *Five questions arising* (pp. 19–22). VI. *Summary* (p. 22). *References* (p. 23).

The rudiments of memory are laid in the perpetual recurrency of the same impressions, and clusters of impressions. The traces which letters, and words, i.e. clusters of letters, leave afford an instance and example of this. And, as in languages, the letters are fewer than the syllables, the syllables than the words, and the words than the sentences, so the single sensory impressions and the small clusters of them, are comparatively few in respect of the large clusters; and, being so, they must recur more frequently, so as the sooner to beget those traces which I call the rudiments or elements of memory. When these traces or ideas begin to recur frequently, this also contributes to fix them, and their order, in the memory, in the same manner as the frequent impressions of the objects themselves.

DAVID HARTLEY, *Observations on Man*, 1749

I. INTRODUCTION

The conception that, underlying the power of recalling past experiences, there exist *traces* in the mind or brain is not only a very old one, but by its simplicity gained wide acceptance among philosophers and psychologists until quite recently.² The idea that, as a necessary condition of reproduction, some *simulacrum* of the past occasion must be preserved carried with it the implication that the alterations in the matter of recall must be quantitative, comparable to the gradual erasure of an inscription on stone, or to the fading of a photograph. Thus many important features of recall were disregarded, though from a theoretical point of view a number of attempts were made to cope with the difficulties of 'fixed, lifeless traces'.³ Many, however (especially physiologists and neurologists), who had occasion to make use of the conception showed themselves notably blind to the explanatory pitfalls entailed.⁴

The experimental work of Bartlett (1932), following that of Phillippe, disclosed a number of features in the process of recall which ill accorded with the doctrine of fixed, lifeless and independent traces. With some of these difficulties the Gestalt school, especially Koffka (1935), attempted to deal by a relaxation of the original doctrine of traces, in the sense of postulating dynamic interaction between them. Within certain relatively simple and delimited fields this extension of the doctrine may be said to have been heuristically valuable. But intolerable complexities and uncertainties arise when the attempt is made to extend the theory generally, as students of Koffka (1935) will note. Bartlett, seeking a fresh approach based upon the theoretical views of Head (1921),⁵ interpreted his

¹ Read at meetings of the Northern Branch, British Psychological Society, 16 May 1953, and of the Bristol University Psychological Society, 8 May 1953.

² Gomulicki (1953) gives a useful account of the history of trace theories.

³ For example, the 'apperceptive mass' of Wundt, and the 'dispositions' of Stout.

⁴ Cf., for example, Munk (1890).

⁵ For reviews of Bartlett's theory and its relation to that of Head, see Oldfield & Zangwill (1942–3), and Northway (1940). Brain (1950) has considered some wider applications of the theory.

experimental findings in terms of a conception which he named the *schema*. This assuredly raises its own problems, but dispenses with the trace as the basic element in the mechanism of remembering.

The trace is a theoretical conception whose concrete exemplar may be any one of those more or less permanent modifications of material substance which, in some more or less direct sense, represent an object or occasion, and which play so large a part in our everyday existence. For all its limitations as a model, it is simple and, as a conception, easily handled in thought. If, however, it must be rejected on account of demonstrated impotence to order empirical findings, it may well be asked what shall replace it. Now it may be noted, that the corresponding method of storing information in terms of permanent traces is equally out of favour with designers of modern computing machines, although for different reasons. In this latter case, a quite novel principle of storage has been evolved. And suggestions have been often made that such a device might form a better model for human memory than does the older one based upon the photograph or gramophone record. Some colour is lent to proposals of this kind by the demonstration of structures in the central nervous system which incorporate inter-connexions similar in kind to those required in the memory-circuits of a modern computing machine. It should be noted, however, that the *known* functions of these structures relate to the temporal prolongation of inhibitory and excitatory effects. No direct connexion with the mechanism of memory as such has been demonstrated. It has been urged, too, that there are grave objections to such a view of the physiological basis of memory (e.g. Eccles, 1953).

The suggestion that the memory-circuit of a computing machine might afford a better analogy for the consideration of remembering has generally been made upon the basis that such a device possesses greater functional flexibility than does the semi-permanent trace. So far as I am aware, however, no effort has ever been made to work out the conception in more detail. In particular, it has not been asked whether a memory founded on this kind of basic element could go any further than does the trace in providing a theoretical formulation agreeable to empirical findings of the kind expounded by Bartlett and his followers. To make this effort is the object of this paper, but discussion will be confined to certain specific points, and no attempt will be made to formulate a general theory of remembering in these terms.

I will first very briefly recapitulate the relevant points of Bartlett's findings; and try to indicate the vital features of his theory. Secondly, the principle of the 'memory-circuit' as used in modern computing machines will be stated, and illustrated by a simple example. We shall then consider whether the latter can be regarded as a suitable element out of which a memory-mechanism could be built, and which would show at least some of the properties required by the known facts.

II. THE CHARACTER OF RECALL

Bartlett, convinced that the experimental constraints employed by Ebbinghaus in the interests of simplicity and quantification in fact prevented the emergence of the salient phenomena of recall, studied the recall of everyday material in the form of pictures and prose. The conditions in which the subject became acquainted with this material approximated to those of ordinary life. His conclusions were, briefly, as follows:

(1) Only in exceptional instances is recall *literal*. In general, changes of various kinds

are introduced into the material. One exception, in which recall may be literal, in effect exactly reproducing the original, is that of word-perfect recitation. Another, not mentioned by Bartlett, is that of immediate reproduction of a quantity of material which falls within the immediate memory span of the subject. In all other cases change is the rule.

(2) Such changes include not only omissions but also qualitative alterations, and even importations. Moreover, these changes are not arbitrary, but occur in accordance with the general principle that the material is converted into a form more in keeping with the cognitive conventions, logical and causal, which are current in the social *milieu* to which the individual belongs. There is *conventionalization* and *rationalization*. As Bartlett views it, the function of this kind of change is to render the material more easily handled in recall.

(3) There is a change in the relative emphasis of different parts of the material. Some elements, particularly details such as names, or other features which in the original are outstanding to the individual and attract his interest, possess especial survival-value. They may, indeed, be elaborated or exaggerated. These dominant details may be preserved in the form of a specific image, and they are thought by Bartlett to act, so to speak, as *labels* which identify the original experience, and as *starting-points* around which the rest of it may be rebuilt when the individual is faced with the need to recall.

(4) Recall is thus essentially *reconstructive*, not *reproductive*. Most of our experience is not retained as such, in explicit or concrete form, but is assimilated into the *schemata*, contributing its quota to the remoulding of their structure and adaptive capacity. Some fragments of experience, on the other hand, survive in a relatively unchanged and explicit form. Recall consists in reconstruction, based upon these latter, by the operation of the *schemata*, which themselves incorporate general laws and principles expressing the uniformities of experience.

Such a view, it is clear, suggests many special problems. But it is at once evident that, if we seek to interpret it in accordance with a mechanical analogy (and the trace-theory itself is a form of mechanical analogy), one important question will arise. How can information be stored in some general or abstracted form, and still allow some explicit, if incomplete, reconstruction of the original? It is sometimes suggested that *no* form of mechanical device could achieve this. But before this can be accepted more detailed consideration of the question is desirable. Accordingly, we shall state and illustrate the general principle according to which the memory-circuit of a modern computing machine operates. We shall then consider the extent to which a system built of such elements could be expected to store information in a generalized form, and subsequently reconstruct the original messages fed into it. It should perhaps be emphasized that such an attempt, though made in specific terms, can do no more than elucidate the basic requirements of the problem. It can, in itself, offer no direct suggestions as to the actual mechanisms, physiological or otherwise, which are involved in memory.

III. CIRCUITAL STORAGE DEVICES

The general principle of the storage devices used in modern computing machines is most simply stated by describing an actual instance. The first requirement is that the information to be stored should be suitably coded. We may, without loss of generality,¹ suppose

¹ It can be shown that anything that is 'information', in the sense that it can be explicitly formulated and represented in symbolic terms, can always be re-coded completely in this form.

that all messages consist of sequences of the scale-of-two digits 0 and 1. In this form the message can be propagated in an electrical circuit as a sequence of impulses. So, for instance, the message 101100011011 will appear as the wave shown in Fig. 1.

Now consider Fig. 2. AB is a tube filled with mercury. At the end A is a transducer element which converts electrical into mechanical changes. These are propagated down the mercury and reach B . Here is situated a receptor element which converts them back into electrical changes. The electrical message-sequence from B can now pass along either or both of two routes. In the first it passes to, and may be used in, other parts of the apparatus. In the second it passes back to A , to be retransmitted down the tube. Deterioration in the form of the impulses, occasioned by the transmission and conversions, is corrected in the unit X , which also makes good incidental energy losses in the system.

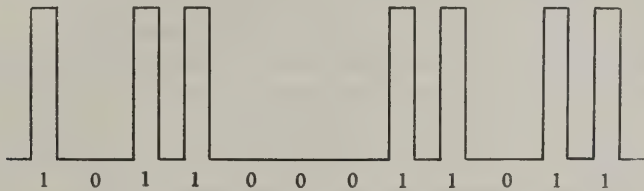


Fig. 1. Pulse-coded version of message 101100011011.

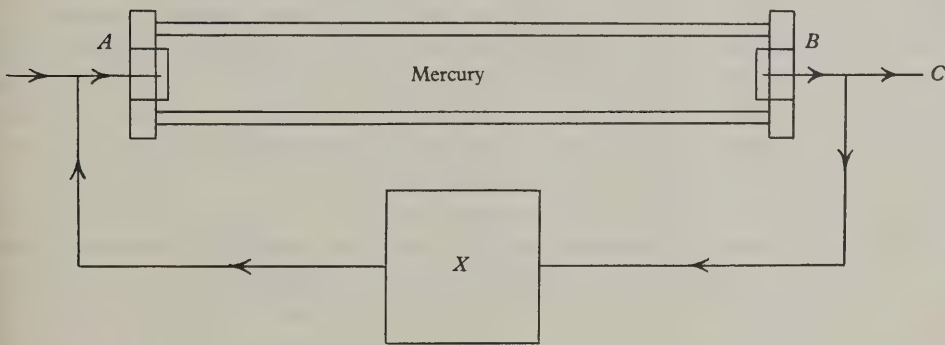


Fig. 2. Mercury circuitual storage element.

If, now, we pass into the system a given message, it will circulate indefinitely, and is effectively stored. If the electrical circuit BXA is broken for a period corresponding to the time taken for the message to pass down the tube, the system is completely 'cleared', and can be used again to store another message. If the information contained in the message is to be made use of elsewhere in the apparatus, it can be passed out along BC , and can be used any number of times in this way. The length of the message which can thus be stored is approximately equal to the number of digits in a sequence which occupies the time taken for one impulse to pass down the tube. The time of transmission in the electrical parts of the circuit may be regarded as negligible.

The general principle of storage is thus a novel one and possesses several outstanding features. Among these are (1) the ease with which the message can be constantly repaired, so that it can be preserved unchanged over long periods of time; (2) the ease with which it can be annulled and wholly disposed of when it is no longer required; and (3) the ease with which it can be made available to other parts of the apparatus when required without

affecting its continued storage. The particular device just described is but one of many possible ones operating upon the same circuitual principle. It happens to be a relatively poor version so far as storage capacity is concerned, and although others have been developed which are much more favourable in this respect, economy in such devices is of basic importance, and is not irrelevant to our further discussion.

IV. MEMORY SYSTEMS BASED UPON CIRCUITAL STORAGE ELEMENTS

So far our discussion offers nothing novel as an analogue for the mechanism of recall. For if we merely postulate a number of independent circuitual storage-elements, we have only returned to another, if perhaps more entertaining and recondite, form of the hypothesis of fixed lifeless traces. The question to be discussed is whether such storage elements could be imagined as built into a system in which there would be some economy of storage without loss of the power to reconstitute the original messages. It does not seem impossible to see, broadly, how this might be done. Consideration of this question raises a number of further points, admittedly speculative, which are of some interest in wider connexions.

We have a 'black box' capable of taking in messages in the form of sequences of 0's and 1's, and storing them. At some future time, on receipt of an appropriate stimulus, the box is to redeliver the original message. For a start we shall suppose that it is to redeliver it in exactly the original form—though we shall later consider relaxation of this condition. The question of what is meant by an 'appropriate' stimulus, and of how it operates to pick out the correct message stored in the box, is, of course, one of great interest and importance. But we shall not, in the first instance, try to answer it. All we wish to do at this stage is to demonstrate the possibilities of economical storage and of reconstruction.

It seems fairly clear that unless the various messages received by the box incorporate *some* detectable elements of common pattern, there is no possibility of economy in their storage. If, however, *any* common patterns run through them, there arises the possibility of re-coding them in briefer form, and so of storing more messages in a given storage space. *One* simple way in which community of pattern might manifest itself, which we may take as an example, is the following. Suppose that all messages handled by the box consist of thirty-six digits, and that the following is one of these:

100100011011100011100100011100100011.

On inspection, it can be seen that if broken up into sub-sequences of three digits each, all of them are of two kinds only, namely 100 and 011:

100.100.011.011.100.011.100.100.011.100.100.011.

If we make 1 stand for 100, and 0 stand for 011, we can re-code the message completely as follows:

110010110110.

In fact, in this particular case, further re-coding is possible in sequences of three, the only patterns in this case being 110 and 010, and the second re-coded version, if $110=1$ and $010=0$ is

1011.

Neglecting, however, for the moment, this possibility of a second re-coding, we may consider how the first re-coding can introduce economy of storage. Suppose that we store

the sub-sequences 100 and 011 in two storage elements, and the re-coded version in a third. The original message can now be reproduced by causing this latter element to draw out the sub-sequences in appropriate order as given by the sequence of digits in the re-coded version. Thus, when the output from the third tube is 1, the sequence 100 is fed out of the box, and when the output of the third tube is 0, 011. The original message can, in principle, therefore, be reconstituted. Now the re-coded message is only a third as long as the original, and requires only a third of the storage space. If, therefore, one storage element was just able to store one *original* message, we can store three re-coded versions in the same space. For nine messages, all susceptible of re-coding on the same basis, only five storage elements are required—three for the re-coded versions, and two for the sub-sequences, in place of nine. It is clear that the advantage will be the greater, the greater the amount of common pattern manifested, and the larger the number of messages possessing this community.

Before we proceed any further a few subsidiary points should be cleared up. The digits 0 and 1 only allow for the re-coding of two types of sub-sequence. If there are more than two, scale-of-two numbers containing more than one digit will have to be used. Thus, if the sub-sequences 1001, 0101, and 0010 are found to constitute all messages, code numbers of two digits will have to be used for each, for instance 10, 11 and 01. Hence twice as much storage space will be required for the re-coded versions as in the case where only two types are present. Secondly, if a number of re-coded versions are to be stored in a single tube, some indication must be provided to serve as a full stop at the end of each message. This can be any sequence not used as a code for any of the sub-sequences. This circumstance, again, makes some slight extra demand on storage space.

On assumptions such as these it is possible to derive an expression for the economy of storage. This would, however, only be valid for the rather special type of pattern assumed in the incoming messages, and its actual form would be of less interest than is the general principle that economy can be effected without impairing accuracy of reproduction. A number of further questions now require consideration.

V. FIVE QUESTIONS ARISING

(1) As we have already remarked, re-coding with consequent economy is only possible if there are some recurrences, uniformities or common patterns in the incoming messages. This is the case in the information from the environment received by the human organism. On the other hand, it is not, of course, suggested that the type of pattern is that very simple one suggested above. By comparison with this latter, the patterns of human experience are complex and subtle. They may only reveal themselves when some amount of abstraction and elimination of irrelevant detail has been carried out, which would correspond with a re-coding process in our analogy. Thus, for instance, the messages 00111011101100110011101 and 110001001110110001 do not present any obvious similarity of pattern until both have been suitably re-coded, by putting 1001 = 1 and 1101 = 0 in the first message, and 110 = 1 and 001 = 0 in the second, when the same re-coded version, 100110, is obtained.

(2) How would the performance of such a mechanism as we have considered compare with the use of Bartlett's schemata in recall? Suppose that the box has had considerable experience of incoming messages, and has acquired a store of type sub-sequences.

These represent common elements in the events in its 'environment'. Treated as a total organized system, which they form in virtue of the various connexions between them which could be generated by re-codings of higher orders, they might be said to form its 'schemata'. For it is upon the basis of them that reconstruction of particular past messages is possible. Furthermore, we may suppose that, once acquired, the system is used to filter fresh incoming messages in accordance with the hierarchical organization which it possesses. Thereby each new message could be re-coded in such a way that only a certain minimum of storage space is occupied. This minimum is that needed to represent the individual features of the message which cannot be subsumed in more general terms derived from the past experience of the box. Some messages will be banal, in the sense that they incorporate little that cannot be subsumed under the schemata. These will demand little storage space, and will induce little reorganization of the schemata. Viewed in terms of the schemata they may be said to be highly probable, and to contain but little information. Other messages, on the other hand, will present novelty, be improbable, and contain much information. These will only to a small degree be capable of reduction and subsumption by the prevailing schemata. They will demand relatively large storage space. On the other hand, they may induce the operation of fresh coding principles, and thus lead to reorganization of the schemata. In this sense it might be said that the box's schemata form an active organization of past experiences. The demands of novel messages of high information-content might be supposed to maintain within the schemata persistently active searching for fresh ways of reorganizing the existing system by exploration of the various alternative possibilities of re-coding. In this way complex, clumsy ways of storing information might, sometimes suddenly, be replaced by much simpler and more economical modes of coding and handling the same information. Perhaps in somewhat similar fashion 'insight' may bring about widespread reorganization of a human individual's response-tendencies.

(3) We have spoken, in this last paragraph, of two supposed processes which require some further comment. These are first the 'filtering' of fresh incoming messages through storage elements which contain type-subsequences, and secondly, a spontaneous activity of searching for fresh methods of re-coding material already stored, whether type-sequences or abbreviated, i.e. re-coded, versions of messages.

The filtering process is, in principle, not difficult to conceive, since type-sequences already stored are available for comparison with fresh material, without jeopardizing their continued storage. To take a very simple instance, suppose that the box contains only two stored type-sequences, 010 and 110, and that a fresh message,

010010110010110110,

is received and stored temporarily as it stands. The output from this storage element is fed to two discriminator units, to one of which is also fed 010, and to the other 110. When the two messages fed to a discriminator unit are identical it responds by sending out a characteristic signal, say 1 for the 010 unit, and 0 for the 110 unit. If the messages are not identical, then the unit does not respond. A sequence of digits is therefore formed by the combined output signals of the two discriminators, in this case 110100, is stored in a further storage element and constitutes the re-coded version. Once the re-coding is complete, the units used for the temporary storage of the original message are cleared and made available for use again.

As to the spontaneous activity of searching for fresh methods of re-coding material already stored, this might be conceived in the following terms. Given a sufficient number of spare storage units for holding temporarily the results of its operation, there is no particular difficulty in supposing the mechanism to explore all the possibilities of re-coding, since this is a type of operation familiar in computing machines. But the mechanism must be controlled in such activities by some general principle, and it must have some criterion according to which it accepts or rejects the results of operations, and continues or ceases its searching activities. Such a general principle might be that of tending to reduce the total amount of storage space occupied by the material stored to date. If this were reduced by any set of re-coding operations, the results would be accepted and stored; and older, now redundant, material annulled. Further activity along these lines would continue until no further economy of storage was produced.

(4) We have suggested that re-coding activities are dependent upon the possession of stored type-sequences, and it might well be asked how, since these must themselves be derived from messages received by the machine, the latter 'gets a start', so to speak, in laying down its 'schemata'. It may, perhaps, be as well not to broach the discussion of this problem, which bears a considerable resemblance to a more famous, but equally troublesome, psychological puzzle. We may remark, however, that if the memory-mechanism we are considering forms part of a more complex machine which interacts with its environment, it may be reasonably supposed that it is endowed with certain 'innate' simple type-sequences.

(5) So far we have been concerned principally to indicate that memory-mechanisms incorporating circuitual storage elements could, in theory, effect economy in the storage of incoming information, and that such economy would take the form of storing general patterns common to a number of messages separately from the special features peculiar to each message. A somewhat similar principle would seem, to judge by the experimental findings of Bartlett and his followers, to operate in the mechanisms of human memory, but in this case a further notable feature is the imperfection of the process, as judged against the standard of literal recall. Reproduction, as we have seen, is rarely literal, and we should consider to what this feature would correspond in the model memory-mechanism we have outlined.

The 'defects' in the products of human memory consist in omissions, alterations, and importations. These changes take place in accordance with the principles of conventionalization and rationalization. The stage in the whole process at which these changes occur has been the subject of some dispute. Bartlett, it would seem (and also the Gestalt school), thought of them as being brought about chiefly as a result of reorganization of material during the period between the original presentation and the act of reproduction. That the act of reproduction itself introduces changes cannot be denied, but these have often been regarded as taking an accidental form, rather than as representing the activities of the schemata. The possibility that an important and characteristic part of the change occurs at the time of presentation, and during the immediate process of storage, has been closely considered by Gomulicki (1952) in a number of pertinent investigations. His conclusion is that a very considerable amount of 'editing' takes place at this stage, so that what is stored has the character of a *précis* of the original material.

Now consider the situation of the mechanism in the face of a message which does not entirely conform to the patterns already incorporated in the machine's 'schemata'.

We might suppose it to act in one of three ways. In the first place, it might simply not store the offending material at all. Secondly, it might condone small departures from pre-existing patterns. Thirdly, it might store those parts of the message which do not conform 'in clear', while re-coding those parts which do. The first alternative is not of interest to us. The second would involve the possession of a criterion by the sequence-discriminators, according to which they would treat as an instance of a type-sequence any sequence which did not depart from type by more than a certain number of digits. So, for instance, if the criterion was 'not more than one digit wrong', and the existing type-sequences were 11001 and 00010, the message

110010101000010110111100100010

will be re-coded 100110, the offending sequences, namely 01010 and 11011, differing from the types by only one digit each, and being coded and stored as the types themselves. If reproduction of this message is called for, it will, of course, appear in a form different from the original, and the alterations could reasonably be said to be ones of 'conventionalization' inasmuch as the material is brought into a form which conforms to the machine's 'past experience' in general.

The third alternative way of handling non-conformatory parts of the material was, it was suggested, by storing such parts in their original form, in much the same way as users of some codes must sometimes leave 'in clear' unusual words not provided for in the code-book. It would be necessary in this case to enclose the sequence in question in a pair of (coded) parentheses to indicate that this material is in clear, so that in the reproductive process such sequences are run straight off without being decoded. If sequences which cannot be assimilated are treated in this way, some points of interest arise. In the first place, the material in question is preserved unchanged and is not subject to processes of 'conventionalization' which may occur in the processes of coding and decoding. Secondly, such outstanding features could act as 'labels', by which such messages might be identified in the course of the further transactions of the system. It seems possible, for instance, that such a label might be operative in the appropriate selection of the message when there is a demand for reproduction. In these respects it might be said that sequences which cannot be assimilated into the schemata resemble the dominant details and images which, in Bartlett's formulation are carried along relatively unchanged, and seem to afford the starting-point for reconstruction.

VI. SUMMARY

An attempt is made to decide whether the memory-model suggested by circuitual storage devices in modern computing machines¹ is better adapted to meet the empirical data on recall brought to light by Bartlett and others, than is the type of model which postulates storage in the form of *simulacra*, or traces impressed upon a medium. It is pointed out

¹ It has been pointed out to me, and it should be emphasized, that the considerations of this paper are by no means dependent upon the *sequential* aspect of processes of storing information encountered in modern computing machines. A type of machine which stored incoming messages on punched cards, automatically searched through these and punched re-coded versions on fresh cards, rejecting the original cards into a waste-paper basket, would do as well in principle as an exemplar. In fact, this paper might, instead of having its present title, have been called 'Sketch-design for an automatic filing system incorporating a principle of storage-space minimization'. The mechanical details of such a model would, of course, be more difficult to imagine, and for this reason the electrical sequential processes we have considered afford a more helpful picture.

that unless the circuital storage-elements are functionally organized in such a way as to allow of re-coding of the stored messages, they offer in themselves no advance upon the ordinary trace hypothesis. However, the flexibility of the circuital device might be used to achieve systematic re-coding upon the basis of common patterns in the incoming messages. These common patterns would then be stored separately from the particular arrangements of them which constitute different messages. Not only could there be economy of storage space, but the original messages could be reproduced in their original forms. Furthermore, when we consider how such a mechanism might handle messages which contain patterns not already held stored, possible analogies to the processes of rationalization and conventionalization, and to the part played by 'dominant detail', emerge. It is suggested that such a system would have some of the properties of schemata as postulated by Bartlett.

Although such a model is not immediately applicable to laboratory experimentation upon memory processes, it might prove useful in re-orientating our conception of memory-mechanisms over a wide field, including problems of recognition and 'set', of perceptual equivalence, of retroactive inhibition and immediate memory span, and of amnesias and paramnesias.

REFERENCES

- BARTLETT, F. C. (1932). *Remembering*. Cambridge University Press.
- BRAIN, W. R. (1950). The concept of the schema in neurology and psychiatry. In *Perspectives in Neuropsychiatry*, ed. Richter. London.
- ECCLES, J. C. (1953). *The Neurophysiological Basis of Mind*. Oxford University Press.
- GOMULICKI, B. (1952). Recall as an abstractive process. Thesis submitted for the degree of D.Phil., Oxford.
- GOMULICKI, B. (1953). The development and present status of the trace theory of memory. *Brit. J. Psychol. (Monogr. Suppl.)*, no. 29.
- HEAD, H. (1921). *Studies in Neurology*, II. Oxford University Press.
- KOFFKA, K. (1935). *Principles of Gestalt Psychology*. London.
- MUNK, H. (1890). *Über die Functionen der Grosshirnrinde*. 2 Aufl. Berlin.
- NORTHWAY, M. L. (1940). The concept of the 'Schema'. *Brit. J. Psychol.* xxx, 316-25, xxxi, 22-36.
- OLDFIELD, R. C. & ZANGWILL, O. L. (1942-3). Head's concept of the schema and its application in contemporary British psychology. *Brit. J. Psychol.* xxxii, 267-86, xxxiii, 58-64, 113-29, 143-9.

(Manuscript received 18 March 1953)

THE CONTINUOUS REGULATION OF SKILLED RESPONSE BY KINAESTHETIC FEED BACK

BY C. B. GIBBS

Medical Research Council, Applied Psychology Research Unit, Cambridge

I. *Introduction* (pp. 24–25). II. *Apparatus* (pp. 25–27). III. *Procedure* (pp. 27–28). IV. *Results* (pp. 28–31). V. *The Tactile Patient* (pp. 32–34). VI. *Discussion* (pp. 34–38). VII. *Summary* (pp. 38–39). *References* (p. 39).

I. INTRODUCTION

The dominant characteristics of skilled movement provide two main hypotheses of the principle or organization which has evolved to control movement and to make effective use of sense data from peripheral nerve endings.

A skilled movement of short duration may be continuously regulated by sense data which are always being fed back from a limb, so that the organism is actuated by the difference between the present state of the limb in motion and the intended final position of the limb. The principle of control would then be continuous 'error actuation' and there would be a direct analogy with a feed-back servo mechanism.

Alternatively, a skilled movement may be 'triggered off' as a unit in response to an external stimulus, but be independent of current sensory control during the initial phase of movement (Woodworth, 1899). It is evident that this phase cannot be continuously monitored by visual feed back in many cases of rapid movement (cf. Fitts, 1951), and Chernikoff & Taylor (1952) have recently confirmed that kinaesthetic reaction time is little less than that in other modalities. Following an argument of Craik's (1947), they claim that this latency would not permit the arrest, or continuous regulation of movements of short duration by the proprioceptive data which are fed to the higher motor centres. A similar view is expressed by Lashley (1951). On this view the selection of an appropriate response is uniquely determined by 'previous experience' and the nature of the external stimulus.

The distinction between the two forms of organization is fundamental, and may best be illustrated by comparing the control principle of a guided, with a projected missile. There is continuous feed back in the former case, and successive modulations of flight behaviour are possible; but in the latter case only one control decision precedes the triggered event, and there is no feed back to permit modification of its flight. Applied to biological mechanisms, the principle of continuous actuation by error would mean that learned responses could be rapidly modified to suit specific environmental conditions, whereas the theory of impulse discharge implies that the main effect of learning is so to 'calibrate' or condition the organism, that a familiar external stimulus can release a relatively automatic and stereotyped response. The survival of organisms has always depended on a good balance between versatility, and the quick action of a ready made response. It is possible therefore that during evolution a compromise has been effected between the principles although one or the other may be dominant in specific conditions.

From the physiological viewpoint, the concept of continuous regulatory feed back from the proprioceptors, has now attained respectable antiquity, since Bell (1826) first dis-

covered the circle of nerves. Experiments on the transfer of high-grade skill (Gibbs, 1951) show that skilled responses are rapidly adapted to new environmental conditions which necessitate a change in the force or extent of a previously learned response. The principle of continuous error actuation, and the relevance of the terminology and physical models of servo engineering are not, so far, generally accepted. The difficulty is pointed out by Luch (1951): 'the protagonists of feed back have been unable to envisage the precise manner in which proprioceptive information is used in the control of movement'.

The main purpose of the present paper is to present a theory of this kind which must be consistent with the main experimental results concerning skilled movement, and account reasonably for the serious objections which have been made to the notion of feed back. Some contemporary data and views are therefore briefly reviewed, and original experiments are described, which were designed to test some critical points at issue; in particular, it is shown that controlled changes in proprioceptive stimulation lead to significant changes in psychomotor performance, and the amount of positive transfer between two skilled tasks is determined largely by the relative adequacy of the proprioceptive data in the first and second tasks.

It is thought that these results can be reconciled only with the theory that skilled responses are continuously regulated by proprioceptive feed back.

II. APPARATUS

(a) *General task.* The apparatus set the operators a task of manual tracking with velocity control. A course-setting mechanism, with two rotating cams, caused a spot of light to wander in two dimensions across the screen of a cathode-ray oscilloscope. The 'Standard' course varied the velocity of the light continuously in two dimensions. The traverse cam, that moved the light from left to right, could be changed to give a more 'difficult' course. The courses are plotted in Fig. 1 and show the movement of the eyes in degrees per second when fixated on the light. Eighty seconds were required for each revolution of the cams; the course was then repetitive, and a standard period of practice was 11 revolutions, i.e. about 15 min. for all operators.

The subjects sat directly facing the screen of the C.R.O., and were 14 in. from it. They held in both hands at approximately elbow height a joystick that could be deflected in any direction about a universal coupling. Elbow supports were provided. The lever was mounted horizontally, pointing toward the operator.

(b) *In the free-moving conditions,* a hand movement to the right or left, up or down, moved the spot of light in the same plane and direction as the hand. The task was that of compensatory tracking, and the subjects tried to hold the light stationary on a fixed cross-wire and central circle of the C.R.O. against the off-setting action of the course-setting mechanism. Any spot movements were errors, and these deviations were summed separately by four counters, which were driven by the central member of the differentials (Fig. 2) and showed errors of 'high', 'low', 'right' or 'left'. The errors were noted at each revolution of the cams and the counters recorded one unit per 0.05 in. of spot movement. The diagrammatical layout of the gear is shown in Fig. 2.

Displacement/velocity or rate control was used and each separate angular position of the joystick resulted in a different velocity and direction of movement of the spot of light. The length of the joystick and the amplification of signal could be altered so that different

26 Continuous regulation of skilled response by kinaesthetic feed back

'gear ratios' could be obtained between the amount of hand movement and the velocity given to the light. The gear ratio (i.e. stick sensitivity) remained approximately constant over the whole range of joystick movement when the sensitivity was set at any given

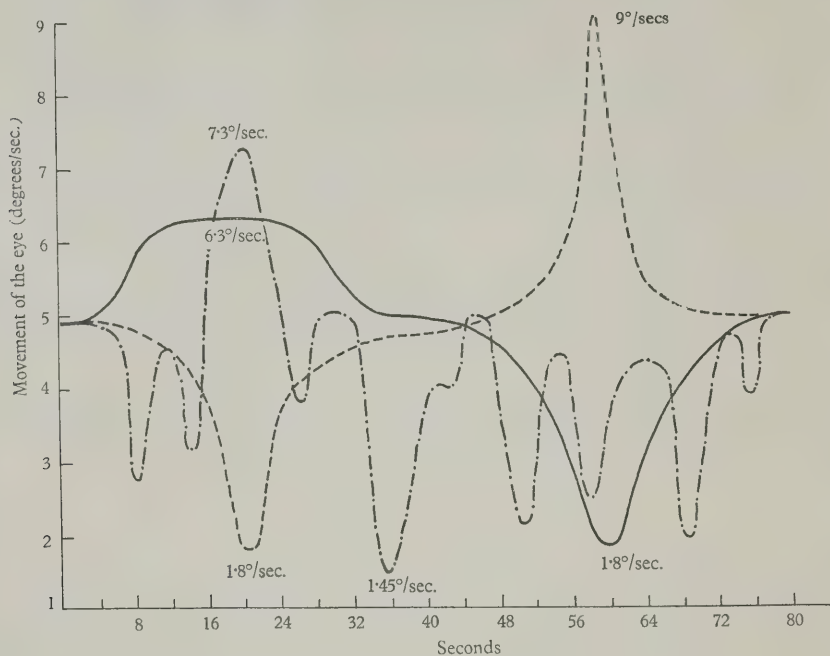


Fig. 1. The courses. Degrees per second at the eye plotted against time. ----, standard elevation course; —, standard traverse course; - · - · -, difficult traverse course.

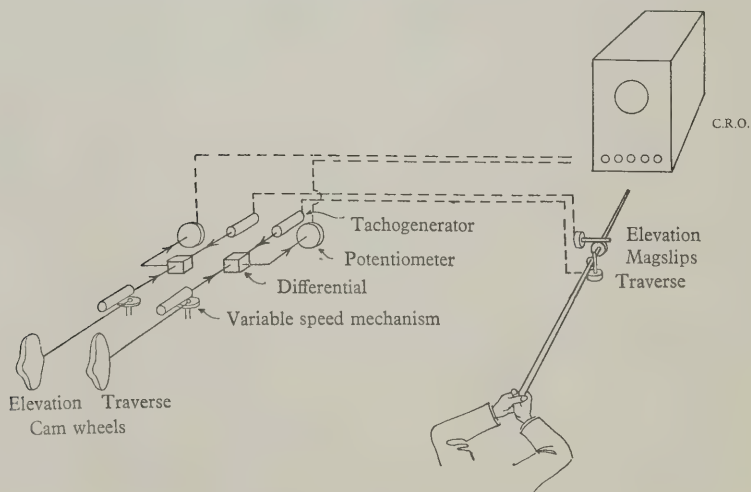


Fig. 2. Diagrammatic layout of apparatus.

value. Several gear ratios were compared to determine the optimal range of this physical variable and a standard ratio of spot velocity per 1 mm. movement of the hand was selected as representing the optimal movement conditions. This standard ratio was used

in comparing the methods of free-moving and pressure control; 1 mm. of hand movement altered the velocity of the spot of light, subtended at the eye, by $0.53^\circ/\text{sec}$. The length of the lever was 19.5 in. from fulcrum to hand grip, and it was counterbalanced.

(c) *In the pressure control arrangement* of the apparatus, the situation was exactly as before except that the joystick was secured at the end remote from the operator and in the position of zero output. Applied pressure would then cause slight deflexions of the lever with spring loading tending to return the lever to centre. The joystick deflexions were converted and amplified, and used to control the movements of the spot, as in the free-moving arrangement. The amplifiers were used in both the pressure and free-moving conditions, but the gain was 17 times greater in pressure control. Variations in applied pressure therefore altered the velocity of the spot, and the pressure/velocity gear ratio was approximately constant over the range of pressure that was used, for each setting of the amplifier.

A number of different pressure/velocity ratios were compared by altering the spring loading and a standard ratio was selected and used in comparing the methods of free-moving and pressure control. In these standard conditions an application of 1 oz. pressure at the lever altered the velocity of the light by $0.26^\circ/\text{sec}$. subtended at the eye. The length of the lever was 19.5 in.

(d) *The time constant of the control system* is an important variable with considerable effects upon human performance. The time constant during the present experiments was 0.08 sec., i.e. there was a time lag between moving the control lever, and the instant when the spot of light reached its new steady speed, and this would be 0.08 sec., if the initial rate of change were maintained.

(e) *Finger operated controls.* The control lever was rotated upwards through an angle of 45° and then operated by the fingers and wrists of both hands, with supported wrists. The main comparisons of the finger controls were made with unchanged gear ratios at the controls ($0.53^\circ/\text{sec}$. per 1 mm., and $0.26^\circ/\text{sec}$. per 1 oz.), since these ratios proved suitable for finger as for forearm controls.

(f) *The single-correction series of experiments* was designed to compare the times taken to correct a given amount of error at the screen, using the standard pressure and standard free-moving methods of control. Small target spots were painted on the screen, 40 mm. distant from the centre. The task was to move the spot of light from the paint spots, to within a small circle of 0.1 in. diameter at the centre of the screen, with the highest possible speed and to a specified accuracy. The course-setting mechanism was running during the tests and the spot tended to wander at a fixed rate, so that the operators had the dual task of (i) making a large corrective movement, and (ii) compensating for a constant course input. Graphic records were taken of the periods between (i) the starting signal to the subject, (ii) the initial movement of the spot of light, and (iii) the instant when the spot of light was arrested within the circle.

IV. PROCEDURE

Subjects. The ninety-five subjects used in the main experiments were Naval ratings aged between 19 and 26 years. About 80% of the subjects had previous experience of visuo-motor skills such as machine operation which were related in some way to the experimental task.

Practice effects. Practice effects were equalized for the test groups by presenting the

conditions in Latin square form. But when only two conditions were compared, two groups of matched subjects were tested after equal training. The subjects were given 80 sec. of preliminary practice on the free-moving control, and the scores were used to divide the subjects into two groups, A and B, which covered the same range of initial ability.

All subjects had a minimum of 2 hr. preliminary practice before any comparisons were made, excepting the thirty-six subjects (groups A and B) who tested initial learning and transfer effects at the two controls.

V. RESULTS

(a) *The comparison of the sensitivities*

Meaningful comparisons could not be made between the controls without preliminary tests to find the optimal conditions for each. Nine subjects therefore tested a range of different sensitivities at the pressure control as shown by Fig. 4A after equal practice on each sensitivity. The same subjects had 3 further days' training on the free-moving control before testing a corresponding range of sensitivities at this control (Fig. 4B). The plotted points show the average total errors for the groups, taken over a full practice period of 15 min.

Accuracy improved nearly linearly at the pressure control as the pressure range was progressively increased, until an optimal point was reached. There was a similar effect at the free-moving control as the rate of movement was increased. After reaching the optimal sensitivity, there was a decrease in tracking accuracy when the pressure and movement range respectively were further increased at the two controls.

Pressure and free-moving gear ratios were selected as the standard conditions on the basis of these results, and they were always used in the subsequent tests of the controls.

There was some small movement at the *pressure control lever* and a total movement of about 0.08 in. was required to vary output velocities between 0° and $9^\circ/\text{sec}$.

(b) *Initial learning and transfer effects*

Thirty-six other subjects were divided into two groups. Groups A and B were matched in initial ability as previously described. An analysis of covariance was performed on the scores obtained in preliminary practice and on the 3rd and 6th day of practice, and it showed that the initial scores had predicted the later scores; thus final performance was dependent on initial ability. The analysis also showed that there were remaining differences between pressure and free-moving control, which could not be attributed to differences in the initial ability of the subjects.

The initial learning and transfer effects for both groups of subjects are shown in Fig. 3 with error plotted against time of practice. The points first plotted on the graphs show the total average errors for each group during the first 80 sec. of practice. The other plotted points show the average group errors per 80 sec. over a complete practice period of 15 min.

Group A first practised the pressure control for 3 days and were consistently better during this period than group B, who first practised the free-moving control. The groups exchanged tasks on the 4th day and the errors of group A were appreciably increased after transfer from the pressure control to the more difficult free-moving control. But in

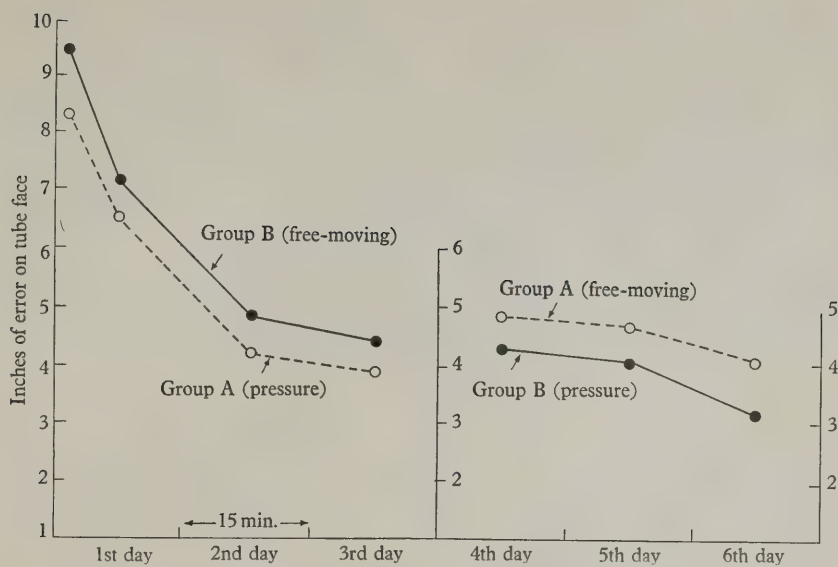


Fig. 3. Learning curves in pressure and free-moving conditions. Group average scores per run. ---, group A; —, group B.

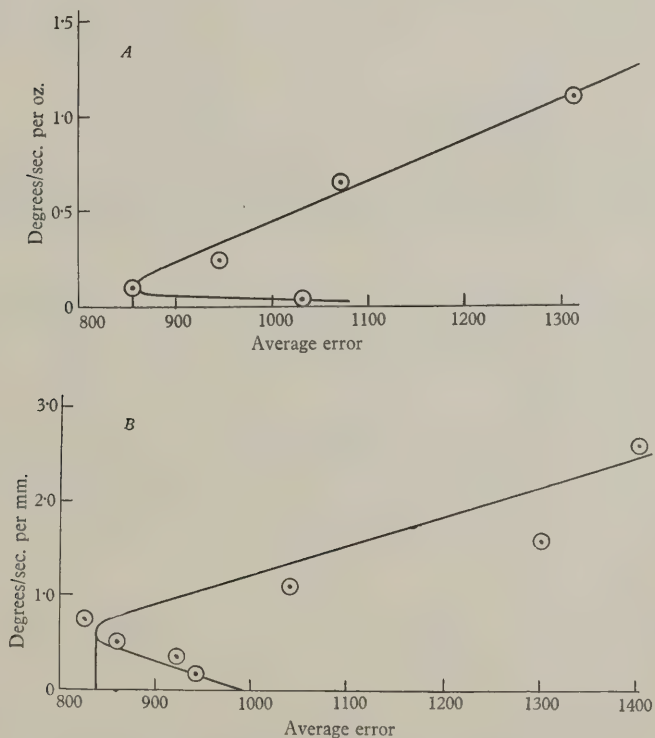


Fig. 4. A, pressure-gear ratios; B, free-moving-gear ratios.

the case of group B there was no deterioration in the scores after transfer from the free-moving to the easier pressure control. There was some disturbance due to the change of task, since the first score of group B after transfer to the pressure control was not equal to the final score of group A, who had learned this control as an initial task. But Fig. 3 shows that this disturbance was partly compensated by the relative ease¹ of the second task. Hence the first score of group B on the pressure control was not inferior to their final score on the free-moving control which they had practised directly. There was appreciable positive transfer from one control to the other in both groups, although the gross external responses were very different at the two controls.

The superiority of the pressure control was maintained after the interchange of groups and tasks, and errors were in the ratio of 3 : 4 in favour of the pressure control on the 6th day of practice. The difference was significant at $P < 0.05$.

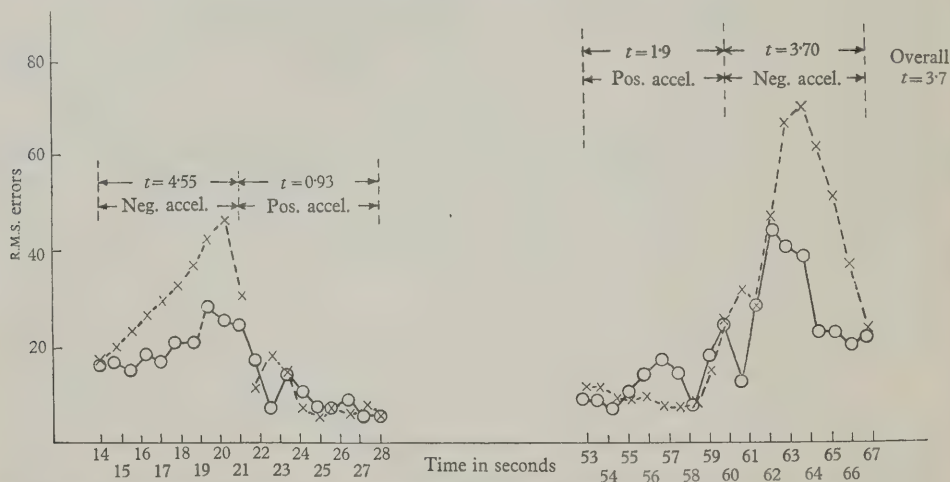


Fig. 5. R.M.S. errors/time. Taken about minimum and maximum rates of elevation course of Fig. 1. $\circ-\circ-\circ$, pressure, mean of five records; $x---x---x$, free-moving, mean of five records. Records from five highly skilled subjects.

Comparisons after longer training. Twenty-four subjects, after longer preliminary practice, then did a further series of trials to compare the two levers on the difficult course of Fig. 1. These tests established the advantage of the pressure control at a high level of probability, $P < 0.001$.

Graphic records of the elevation errors were taken from these subjects at the same time as the counter readings, to make sure that the method of scoring was not determining the answer being obtained. The graphic records from five of the subjects were analysed in detail, and Fig. 5 shows that the pressure control was again superior using the root mean values of the squared errors as the criterion. Although the records were taken from different people, they had striking similarities and showed clearly that the pressure control appeared to greatest advantage when it was relaxed towards centre under the influence of spring loading (i.e. during negative target accelerations) and also at cross-over points in the target course. Large tracking errors developed at these cross-over points since there was a sudden change in the direction of target motion, requiring a corresponding change in the continuous tracking adjustments of the limb.

¹ Definitions of the difficult-easy continuum in skill have previously been given (Gibbs, 1951).

During the 7 sec. of negative acceleration (towards the minimum velocity of the 8°/sec. of Fig. 1) the R.M.S. errors for the pressure control were compared with those for the free-moving control and a 't' value of 4.5 was found ($P < 0.01$). But during the next sec. of positive acceleration away from this speed there was no significant difference between the controls. There were similar effects about the maximum required speed of the course (i.e. 9°/sec. in Fig. 1). Again there was no reliable difference between the controls during the 7 sec. of positive acceleration, but the 7 sec. of negative acceleration following the peak speed showed a great advantage in favour of the pressure control ($t = 3.7$, $P < 0.01$).

Fig. 5 shows that the operators started to diminish the major errors at the cross-over points about 1 sec. sooner with the pressure than with the free-moving control since the peaks of the two error curves have a temporal separation of about 1 sec. at the cross-over points.

(c) *Finger operated controls*

A further group of fourteen subjects were used to compare the controls when operated by the fingers after equal practice in pressure and free-moving conditions.

The pressure control was then markedly better than the free-moving control with the easy course set at the apparatus ($t = 4.6$, $P < 0.01$). The average scores were slightly better when the controls were operated by the fingers as compared with the forearms, but the difference was not statistically reliable, since $t = 0.73$.

The finger-operated pressure and free-moving controls were again compared, with the difficult course set on the apparatus. The superiority of pressure control was again highly significant at $t = 5.85$, and $P < 0.01$.

(d) *The single-correction series experiments*

Sixteen subjects compared the pressure and moving controls when large discrete corrections were made to a high level of accuracy. There are analysis difficulties in reaction and movement times, and a large number of recordings were therefore made, interpreted, and cross-checked by two observers. These tests showed that significantly less time was required to operate the pressure lever than the free-moving lever, after the signal to commence a correction ($t = 3.44$, $P < 0.01$). The time taken to *complete* this correction to a given level of accuracy was also significantly less when the pressure control was used ($t = 6.9$, $P < 0.01$). There was agreement between the 1 sec. difference in completing an accurate single correction and the 1 sec. displacement in the peaks of the two tracking curves of Fig. 5. Major factors in pressure superiority were therefore the reduced delay in modifying a previous muscular state, and settling more rapidly to a new rate of muscular contraction with this control. The results are set out in Table 1.

Table 1.

Condition	No. of subjects	No. of trials	Average time to start (sec.)	95% confidence interval ¹	Average time to complete (sec.)	95% confidence interval
Pressure	16	64	0.313	0.047 to 0.27 sec.	3.073	0.74 to 1.34 sec.
Free moving	16	64	0.426		4.113	

¹ The average difference was 0.113 sec., but this value is subject to sampling errors, whereas there is 95% confidence that the real difference lies between 0.047 and 0.27 sec.; similarly, the real difference in completing correction was 0.74–1.34 sec. The times are rather long compared with some other tests of reaction and movement times, but this is due to the long time-constant of the apparatus, the ambiguity of the display, and the high level of final accuracy which was demanded.

VI. THE TABETIC PATIENT

A tabetic patient, aged 55 years, was tested on the apparatus. He was normally right-handed and usually employed as a general labourer. At the time of the tests the general reflexes were present but diminished, and the plantar response was negative. Sensitivity to pin prick was diminished at both hands and wrists, especially on the left side. The proprioceptors were impaired at the ankles and knees as shown by swaying on the Romberg test, and by diminished appreciation of movement and passive position. The pupils were typical of *tabes dorsalis*, and there were positive serological findings in the blood.

There was some weakness of the extensor muscle of the left wrist and wastage of the left hand. The patient could not adequately regulate the grip of this hand; but it could be used when sweeping with a broom, and for supporting weights such as a bucket of coal.

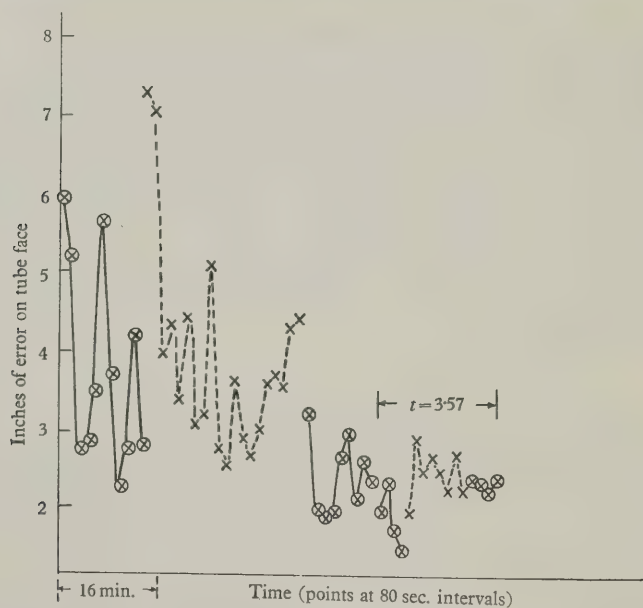


Fig. 6. Errors/time. Elevation record only on counters learning curve of tabetic patient. Fingers on right hand (intact). (Plotted points show total errors per 80 sec. practice.) $\otimes$ — $\otimes$ — $\otimes$, pressure, right hand (intact); $\times$ — $\times$ — $\times$, free-moving, right hand (intact).

The preliminary clinical tests of his ability to discriminate with eyes closed between different directions, rates and ranges of limb movement showed some impairment of proprioceptive sensibility in the finger joints of the right hand. This was slight, however, as compared with fingers of the left hand where there was considerable loss of sensitivity. Impairment was progressively less marked at the elbow and shoulder joints of the left arm.

The subject first practised tracking, using only the fingers and wrist of the intact right hand. The total error at the cathode-ray tube was 6 in. during the first 80 sec., but this was reduced to about 2 in. per 80 sec. at the end of 5 days of continued practice (Fig. 6).

The control was then transferred to the fingers and wrist of the impaired left hand, and there was an immediate and marked deterioration in performance with 13 in. of error at the tube face during the first 80 sec. of practice after transfer (Fig. 7). This was more than

vice the initial error with the intact hand, despite the usual finding of considerable lateral transfer in this kind of skill (Woodworth, 1938).

Seven minutes of practice were given in these conditions, and the placement of the control was again altered so that the fore and upper left arm could now operate the lever.

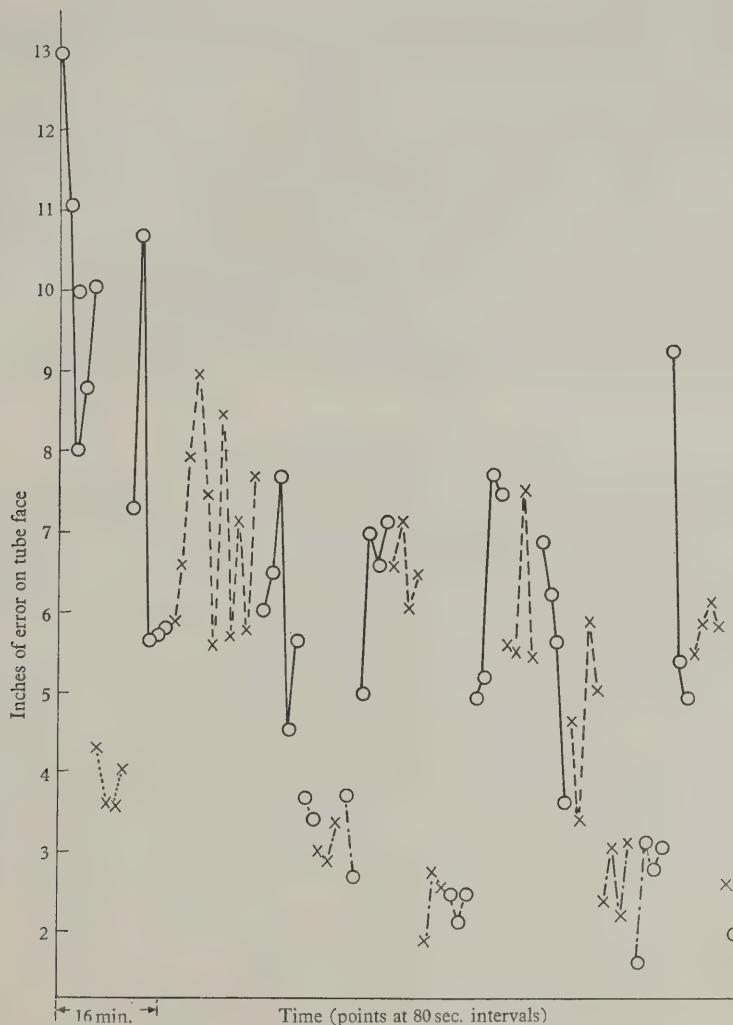


Fig. 7. Errors/time. Elevation record only on counters. Transfer curve to fingers of left hand (lesion) of tabetic patient, with check readings of fingers of right hand (intact) and four readings of transfer to whole left arm (intermediate impairment) (plotted points show total errors per 80 sec. practice). ○—○—○, pressure, left hand (lesion); ×---×---×, free-moving, left hand (lesion); ×---×---×, free-moving, whole left arm (intermediate impairment); ○—○—○—○—○, pressure, right hand (intact); ×---×---×---×, free-moving, right hand (intact).

The clinical tests had indicated intermediate impairment at these joints, and this was confirmed by the tracking scores, since the errors fell to 4 in. of error per 80 sec. after this second transfer. Five minutes' practice was given in the upper arm conditions, and the control was again reset so that practice continued with the impaired left fingers and wrist for a further 6 days. There was some improvement over the first 45 min. of practice with the impaired hand, and the subject learned to make periodic corrections in the correct

sense so that the errors were reduced to about $5\frac{1}{2}$ in. per 80 sec. But the rates of muscular contraction were seldom appropriate. In general, there was abnormal latency in initiating large corrective adjustments, and slow and continuous oscillations of the spot of light about the correct aiming point, a condition of instability which is also characteristic of servo mechanisms in certain conditions of phase lag in their feed back.

Only one tabetic patient was available for the tests, but both the intact and impaired states were found in one individual, thus equating the usual variables of intelligence, motivation, learning, etc.

The main purpose of these tests was to confirm the transfer results which had been found with intact subjects, when the level of proprioceptive stimulation was altered by mechanical means. For the rest, the results confirmed that vision can compensate only to a limited extent for the defects of locomotor ataxia; and the results are consistent with the findings of Sherrington (cf. Fulton, 1949), and also with the demonstrations by Holmes (1922) of disturbances in the direction, range, rate and force of movement following unilateral cerebellar lesion. Such agreement would be expected because of the heavy proprioceptive input to the cerebellum.

VII. DISCUSSION

Pressure superiority is not explained by motor effects at spinal level

The considerable difference in the proprioceptive discharge during isotonic and isometric contraction could reasonably be regarded as the main experimental variable. There was more rapid release and settling to new patterns of contraction with the pressure control (Fig. 5; Table 1) to suggest that the discharge had important central effects, in addition to any smoothing and inhibitory effects at spinal level.

It was, however, important to consider first the possible advantages of pressure control in the motor aspects of performance. The movement range was negligible (0.08 in. maximum), and the time required to make such movements might be very small. This possible advantage could be discounted, however, in a paced task such as tracking where the correction speeds are externally imposed and little is gained by providing a control which can potentially exceed the target schedule. The sensitivity tests (Fig. 4B) showed that tracking errors were in fact increased by reducing the movement range below an optimal value, despite the assumed reduction in movement time. A lever with a small movement range could only be operated effectively when pressure cues were provided to compensate for the lack of movement cues (Fig. 4A, B).

Again, the errors of overshooting or leading a moving target might conceivably be reduced by a spring-loaded lever which opposed the dominant direction of tracking movements. But Fig. 5 shows that the pressure control was superior when spring loading was acting to accelerate, rather than to retard, the limb. Spring loading would tend to alter the force distribution in the muscles (i.e. controlled relaxations might tend to replace contractions in some phases), but the required force change was negligible in either case as compared with the maximum force output which can be developed if necessary in about 0.04 sec. Any time difference between generating and dissipating such force output is not of an order to explain the differences which were in fact found between the controls.

Motor effects, or sensory effects at spinal level therefore do not explain any significant proportion of the results, whereas a complete explanation is possible, if it is postulated

that the sensory discharge is used continuously at central level, to regulate the force output in muscles.

Current hypothesis of the impulse discharge of movement

It is by no means generally believed, however, that the kinaesthetic discharge has important central effects, and the hypothesis of impulse discharge is currently supported by Lashley (1951) and Chernikoff & Taylor (1952). Their main objections to continuous feedback are based upon the somewhat conflicting evidence concerning kinaesthetic reaction time: about 0.125 sec. may elapse between an unexpected kinaesthetic stimulus and the release of an appropriate voluntary response (Craig, 1947; Vince, 1948; Hick, 1949; Chernikoff & Taylor, 1952). Lashley (1951) discusses the real difficulty of reconciling that latency with the observed times of movement, e.g. a single finger may be tapped at the rate of 10 strokes/sec., and the successive finger movement of a skilled pianist can be made in definite order at about 16 strokes a second. He concludes that 'the sensory control of movement seems to be ruled out in such acts'. The higher motor centres must release unmonitored impulses which have no effective sensory correlates if feedback lags are indeed constant at 0.125 sec.

Objections to the impulse hypothesis

The initial assumption might well be challenged, since the reaction times to unexpected stimuli can scarcely be transposed to practised skills or to the monitoring¹ of a movement which has already been released, but impulse theory may best be discounted on more general grounds. The inevitable implications of the theory have been clearly stated by Chernikoff & Taylor (1952): 'the control of movement consists of the perception of error, selecting a response pattern judged on the basis of previous experience to be appropriate to null out the error, and triggering the entire sequence'. But a rapid movement has a continuous phase, which is certainly not monitored by visual cues (cf. Fitts, 1951) and the lag in obtaining visual knowledge of results is particularly marked in such skills as golf, bowling, etc. If kinaesthesia is also to be excluded, then the proposal of impulse theory is that an appropriate response can be learned, and selected at will from a repertoire, without any sensory discrimination of its suitability until the initial phase of movement is completed.

The theory of knowledge of results is essentially a theory of feedback and is not compatible with the notion of impulse discharge. Observations of high-grade skill show that the feedback stimuli which are learned as a basis for control must necessarily precede the visual knowledge of error. The feedback stimuli of visual error tend to disappear as the correct anticipatory cues are learned, and this common observation points to two relatively distinct sources of feedback. It is no purpose of skill to acquire knowledge of visual errors; it is, rather, to find the means to eliminate such errors. It is therefore significant that the spino-cerebellar tracts conduct at the extremely high rates of 140 m. per second, so that only a few milliseconds are taken for conduction from dorsal roots to cerebellum.

Centripetal conduction time is probably significant in the learning of high-grade skills, but the time taken to act on sense data is much longer and is the critical time in the actual

¹ The term monitoring is more specific than feedback. In this context it is defined as a process of continuous comparison between an intended state and the current state: continuous corrections are based upon the discrepancy and act in the sense which is appropriate to reduce the error.

control of movement. Feed-back theory can give a reasonable basis for learning, but it must also suggest a means whereby the kinaesthetic discharge could be used to monitor and terminate the initial phase of a released movement, despite the known transmission lags of the central nervous system.

The characteristics of proprioceptive nerve endings

The characteristics of several receptors have been determined by Matthews (1933) and must be considered in developing a theory of the control of movement. In resting or static conditions, the frequency of discharge in the muscle receptors is proportional to the logarithm of tension on the muscle. However, when the muscle is stretched, that simple relationship is disturbed and there is then only a correlation between the rate of change of frequency and the rate of stretching the nerve ending, but not between the absolute frequency and the absolute amount of stretch at any instant. The peak discharge during stretch, within the limits of the rates which were tested, may be over four times the final resting discharge in primary nerve endings, and more than twice as large in the tendon organs which lie in series with the muscle. Boyd (1953) has demonstrated a similar effect in the discharge from proprioceptors in the knee joint of a cat. The known discharge in isotonic conditions can therefore only define the rates of movement during a rapid movement; the final resting state of position is only defined by the discharge after movement has ended, when the receptors have again settled to equilibrium.

There is no rate of stretch component to complicate the discharge during isometric contraction, and the discharge therefore continuously defines the important control quantity of present state, since the absolute frequency is logarithmically proportional to the absolute applied tension. This discharge could define both the present tension, and rate of change of tension in the muscles.

Lissman (1950) quotes data by Katz, that the discharge in some primary endings is considerably boosted in isometric conditions. There is reason to believe that the isometric discharge is carried in fibres of large diameter whereas the secondary nerve endings, with a low threshold to rate of stretch, have smaller fibres and are inactive during isometric conditions. Isometric sense data may therefore be conducted more rapidly, and may be more adequate in control because of the additional degree of information in this discharge.

The action of cutaneous receptors is not important in the present context, since they have little effect in the control of movement (Sherrington, see Fulton, 1949). Subsidiary experiments by the present writer have shown that the accuracy of tracking with pressure control is not impaired when the controlling fingers are strapped very tightly to the control lever.

A hypothesis of feed-back control

It is perhaps axiomatic that movements of definite extent cannot be made without knowledge of the starting and end position of the limb. Vision or proprioception could adequately define the initial position of a stationary limb, but the quoted evidence (Fitts, 1951) shows that only proprioception need be considered as a possible means of first halting a rapid movement by current sensory control.

A rapid movement cannot be terminated by a system of kinaesthetic 'position' indicators which might conceivably be arranged to indicate that a desired position or angle had been reached. Such movements would be overshoot because of the time taken by feed-back impulses to traverse the higher centres and to return as motor impulses

estimated by Ruch, 1951, as 19 to 36 msec.) plus the time taken by contractions to arrest the moving limb. Positional control is more definitely precluded by the action of the known receptors. A more reasonable principle of current control is that a rapid movement is first halted by integrating, over a period of time, the feed-back signals of the rate of movement, i.e. the extent of movement could be estimated by integrating a known rate of movement over a known period of time.

The principle of the rate control of movement would imply that in the apparently continuous task of tracking, there are only intermittent estimates of the successive positions reached by a moving limb. Such tasks would therefore have definite temporal phases, corresponding to the length of the integration period, and the end of one phase of movement is necessarily the starting-point of the next. There would be a definite limitation in such a system, that a released movement must necessarily run for an integration period to give the initial data for the next phase, since a new movement of definite extent cannot be organized without knowledge of its starting-point.

The hypothesis would completely account for the present results, that, with the pressure control, there was more rapid response to a sudden large error and quicker settling to a change of target behaviour. This was probably because the isometric discharge continuously defines the 'present' state of tension in the limb. The delays which are necessarily involved in any process of integration are presumably reduced by the extra degree of information from the isometric discharge. The extensive general evidence of control intermittency has been reviewed by Welford (1952). It is suggested that the present results can best be regarded as a special case of the psychological refractory phase, and the present hypothesis may therefore account, at least in part, for this general effect in movement.

The principle of rate control clearly disposes of the main objections to feed-back theory, namely that the known transmission latencies of the central nervous system preclude the possibility of current sensory control. The duration of movement (i.e. the integration period) is predetermined to allow for these latencies, and the continuity and rate correlation of the feed-back impulses permits purposive control and monitoring of the rates of movement before the limb has reached its final and intended position.

The hypothesis may also throw light on the reasons for the well-known tendency toward time constancy when making movements of very different extent (cf. Fitts, 1951). The two controlled variables of a velocity system are the rate and the duration of movement. Clearly, the difficulties of learning a repertoire of movements of different extent could be greatly reduced by a tendency to keep one controlled quantity constant and learning only a range of variation in the other. Hence time constancy may well appear as a device for greatly reducing the repertoire of learned items, while still retaining the required versatility.

Craik (1947) has previously suggested that anticipation is essentially a process of extrapolation from current sense data. It is clearly impossible to extrapolate from a present to a future bodily state without current knowledge of the rate of change of state. Hence, the continuity and rate correlation of the kinaesthetic discharge gives a continuous sensory basis for anticipation, as would be required by Craik's hypothesis.

Learning and transfer effects

The hypothesis of velocity control is therefore entirely consistent with physiological findings, the dominant characteristics of skilled movement, and with the theory of knowledge of results. But the consistency of feed-back theory does not necessarily

preclude the alternative notion of ballistic impulse, and, in this connexion, the results of transfer experiments are of decisive importance.

The present experiments were undertaken largely to confirm and explain previous data on the transfer of skill (Gibbs, 1951) which had shown that physical similarity of stimuli and responses in a first and second task was not a particularly meaningful element in the transfer of high-grade skill. Of greater importance were the optimal zones in a range of exteroceptive or interoceptive sense data (e.g. Fig. 4 A, B), and the relative adequacy of the sense data in the first and second task was an important determinant of the amount of transfer between them. These earlier findings are confirmed by the results of the present experiments since there was no deterioration in group performance after transfer from the free-moving to the pressure control (i.e. from meagre toward adequate sense data), whereas there was a sharp decline in performance after transfer with the opposite order of presentation. Again there was little transfer of skill from the intact to the impaired hand of the tabetic patient (Figs. 6 and 7). The previous experiments also showed high positive transfer between two different tracking courses, which presented entirely different problems in space and time, indeed, the relative content or difficulty of such tasks could be arranged so that indirect was equal to direct practice.

Such results show clearly that the organism is not 'calibrated' by learning. Stimuli and responses are not learned as a record of events in space and time, nor do specific external stimuli trigger a blind series of stereotyped muscular contractions, which run their course independently of current sensory control. Skilled behaviour can be rapidly adapted to different conditions in space and time, but the greatest transfer is obtained when there are equally adequate kinaesthetic data in the second task to define the changed conditions at the periphery. These data are fed back continuously and define the departure from a desired condition, so that the organism is in fact continuously actuated by error during a skilled movement.

It seems, therefore, that a main purpose of learning a skill is to filter and to interpret the anticipatory cues of the kinaesthetic discharge. The periodic oscillatory errors in the early stages of learning a tracking task (e.g. riding a bicycle) are well known, and are probably due to dependence on visual feed back, with consequent delays in knowledge of results. Oscillatory errors were very marked in the records of the tabetic patient throughout his training, and the stability of skilled performance can easily be disturbed by the use of high sensitivity controls (Fig. 4 A, B), which have inadequate movement or pressure cues.

The problems of learning, in relation to high-grade skill, may therefore hinge upon relatively small time differences in the feed back from vision and kinaesthesia, and upon the anticipatory characteristics of the kinaesthetic discharge. The learning and sensitivity experiments, and the sluggish responses of the tabetic patient do not support Lashley's (1951) suggestion, that the rapidity of skilled movements precludes the possibility of current sensory control. It seems much more probable that the exact converse is true; the need for knowledge of results from kinaesthesia sets a real limit to the rates of successive innervation in synergic muscle groups.

VII. SUMMARY

Ninety-five normal subjects and one tabetic patient were tested in experiments which were designed to alter systematically the response in a visuo motor skill and, thereby, to alter the proprioceptive discharge in the controlling muscles.

Two types of joystick were used to control the output of a velocity control servo mechanism in a tracking task. One lever was of the normal free-moving type, and there the muscular contractions were isotonic when the subject was making controlled responses by displacing the lever. The other had strong spring centring, but it could be deflected slightly by applying pressure, so that the controlled contractions of muscle were virtually isometric. Changes of limb position, and changes in the force exerted by the limb could therefore be compared, using the criteria of speed and accuracy.

Isometric contraction with the pressure control was definitely better when making both discrete and continuous responses, learning was easier, and differential transfer effects appeared which also favoured the pressure control. It is thought that this superiority is due to the marked difference in the proprioceptive discharge in the two contractile states. The immediate implications are that skilled movements are continuously regulated by incoming kinaesthetic data which are generated by the limb in motion and are continuously fed back to the higher motor centres. The broader implications are briefly discussed, especially how cause and effect operate in sensory organization: a theory of control is therefore developed.

My thanks are due to Dr W. E. Hick and Dr N. H. Mackworth for reading and discussing the paper, and to Mr N. L. Webb for advice on the statistical analysis. Mr D. Russell Davis kindly made available the help of the tabetic patient and his clinical data.

REFERENCES

- ALL, C. (1826). On the nervous circle which connects the voluntary muscles with the brain. *Phil. Trans.* pt. 2, 163-75.
- BOYD, I. A. (1953). Nerve impulses from proprioceptors in the knee joint of a cat. *J. Physiol.* **CXIX**, 8/P.
- CHERNIKOFF, R. & TAYLOR, F. V. (1952). Reaction time to kinaesthetic stimulation resulting from sudden arm displacements. *J. Exp. Psychol.* **XLIII**, 1-8.
- CRAIK, K. J. W. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.* **XXXVIII**, 56-61; 1948, **XXXVIII**, 142-8.
- FRITS, P. M. (1951). Engineering Psychology and Equipment Design. A chapter in the *Handbook of Experimental Psychology*, pp. 1328-9. Ed. S. S. Stevens. New York: J. Wiley and Sons.
- GILTON, J. F. (1949). *Physiology of the Nervous System*, pp. 168 and 518. New York and Oxford.
- GIBBS, C. B. (1951). Transfer of training and skill assumptions in tracking. *Quart. J. Exp. Psychol.* **III**, 99-110.
- HICK, W. E. (1949). Reaction time for the amendment of a response. *Quart. J. Exp. Psychol.* **I**, 175.
- JOLMES, G. (1922). The Croonian lectures on the clinical symptoms of cerebellar disease and their interpretation. *Lancet*, 202, 1177-82; 1231-7.
- ASHLEY, K. S. (1951). The problem of serial order in behaviour. A chapter in *Cerebral Mechanisms in Behaviour*, ed. L. A. Jeffries. New York: J. Wiley and Sons.
- ROSSMAN, H. W. (1950). Proprioception. A chapter in *Physiological Mechanisms of Animal Behaviour*. Soc. Exp. Biol. Symp. no. iv. Cambridge University Press.
- MATTHEWS, B. H. C. (1933). Nerve endings in mammalian muscle. *J. Physiol.* **LXXVIII**, 1-53.
- WELCH, T. C. (1951). Motor Systems. A chapter in the *Handbook of Experimental Psychology*, pp. 199 and 204. Ed. S. S. Stevens. New York: J. Wiley and Sons.
- BRINCE, M. A. (1948). Corrective movements in a pursuit task. *Quart. J. Exp. Psychol.* **I**, 85-103.
- WELFORD, A. T. (1952). The psychological refractory period and the timing of high-speed performance. *Brit. J. Psychol.* **XLIII**, 2-19.
- WOODWORTH, R. S. (1899). The accuracy of voluntary movement. *Psychol. Rev. Monogr. Suppl.* **III**, 13.
- WOODWORTH, R. S. (1938). *Experimental Psychology*, pp. 181-7. London: Methuen and Co.

(Manuscript received 15 April 1953)

THE ADEQUACY OF EVERYDAY QUANTITATIVE EXPRESSIONS AS MEASUREMENTS OF QUALITIES

BY D. SHEPPARD

National Institute for Research in Dairying, University of Reading

Quantitative terms, such as 'Very' or 'Rather', are frequently used for describing qualities, both in everyday conversation, and by psychologists for certain types of rating scales. Admittedly such quantitative terms may be imprecise in meaning, yet it is generally considered that there will be fair agreement in the interpretation of these terms, provided that the quality which is being described is well known.

Psychologists have shown that judges may disagree markedly when they have to assess independently the qualities of any one thing, but it is not easy to tell from such evidence whether the judges disagree about the merits of the thing being judged, or whether they merely have different opinions about the meanings of the words with which they choose to describe it. In addition, it is not easy to find qualities that large numbers of people can judge expertly, and for which certain people will not have special preferences or biases, as for example with assessments on flavour.

For certain purposes, however, it is necessary to obtain absolute judgements for qualities of materials in terms of such quantitative descriptions. Thus when obtaining consumer ratings for food qualities at the National Institute for Research in Dairying it was an easier task for the raters to use everyday quantitative terms to describe the values, than to allocate marks for these qualities by a points system.

A set of descriptions so obtained may be sufficient for some purposes, but in most cases it is desirable, if theoretically unjustifiable, to be able to allot numerical values to these descriptions, so that at least some simple statistical treatments can be used on them. The most usual procedure seems to be to allot serial numbers to the quantitative descriptions, and then to treat them as if they were equally spaced one from another, and this, although an arbitrary procedure, often leads to fruitful results.

In view of these considerations, some simple experiments were carried out in an attempt to find out how people agreed about the use of quality descriptions for items of everyday experience, and also to attempt to discover a better method of allotting numbers to these quality descriptions, for the purposes of statistical treatment.

METHOD

For very many reasons, it is hard to compare people's subjective scales by asking them how they would describe given samples of material. Thus, for example, there are known effects which make such a comparison difficult. Amongst these are the 'central tendency' (Hollingworth, 1910), the fact that ratings tend to be a function of the stimulus range presented (Tresselt, 1948; Sheppard, 1953) and other difficulties that are incurred for the reason that objectively equal intervals are unlikely to be subjectively equal.

An unusual method of approach, which avoids these particular difficulties, is to present the subject with a list of quality descriptions, and ask him to say what objective measure

the quality concerned corresponds to them. This can of course only be done when the subjects are familiar with the objective measurement concerned and are used to thinking in terms of it, and provided that the qualities 'form a continuum'.

For this purpose then, a number of people were asked the following questions:

Test A. How far is (a) very far, (b) rather far, (c) very near, (d) near, (e) far, (f) rather near? (g) What distance lies in between, and is neither far nor near?

These questions were asked in this order, and the experimenter noted down the answers.

Test B. If you were going for a walk somewhere, what distance would you regard as being 'very far', etc. Please fill in what you consider to be suitable distances into this rating scale.

Here the descriptive qualities were the same as in test A except that 'not very near' and 'not very far' were added. The 'positive' values were written out in a definite order against a horizontal line. The subjects read the descriptions and filled in the appropriate numerical values themselves in this test, and also in tests C and D.

Test C. If you were waiting for a bus, what sort of time would you have to wait for you to describe as being 'a very long time', etc. Please fill in the appropriate times into this rating scale.

Here the method was the same as for test B, with the qualities as follows:

(a) a very long time, (b) a long time, (c) a rather long time, (d) neither a long nor a short time, (e) a rather short time, (f) a short time, (g) a very short time, (h) a not very long time, (i) a not very short time.

Test D. How much would a single man have to earn, in a week, for you to describe him as being 'very rich', etc. Please fill in the appropriate sums into this rating scale.

Here the qualities were similar to those in test C, but ranging from 'very poor' to 'very rich'.

Test E. In addition, some 300 subjects were asked to fill in descriptive values onto a 7 in. straight line, the written instructions for this being as follows:

You must often have made remarks such as 'that is very good' when describing the quality of different things. Do you consider that, for example, 'good' is as much better than 'average' as it is worse than 'very good'?

A quality scale ranging from a point representing 'extremely good' at one end, to a point representing 'extremely bad' at the other, is given below. Suppose that equal distances along the line represent equal increases, or decreases, in quality, can you indicate where *you* think the following quality judgements could occur in this scale? Arrange the quality descriptions so that they are not only in what you think the right order, but so that the *distances between* each quality description are what you consider to be correct.

Very good, Rather good, Good.

Very bad, Bad, Rather bad.

Average.

Not very good, Not very bad.

Use a ruler if you wish, and mark in the appropriate positions for each of these qualities by a stroke crossing the line, and name the intersection.

Please do not discuss this with anyone else, until you (and they) have completed the task.

Note. (1) The same point on the scale may be used for more than one quality description if you wish. (2) Certain of these descriptive terms may differ in meaning if one or other of their constituent words is emphasized. Try to consider them, therefore, as if no one word in each term were emphasized more than any other.

For certain of these forms the word 'Moderate' was substituted for 'Average' and the word 'Fairly' for 'Rather'. In other forms 'Quite', or 'Fairly', were put in place of 'Rather'.

A large number of subjects answered these questions, although different sized groups were used in each test. Thus only ten subjects did test A, eighty-five subjects did tests B, C and D, and approximately 300 subjects did test E. Test A was, in fact, only a preliminary one, and the largest number of subjects were used for test E since the results of this were of practical service to the writer.

Careful selection of subjects was not considered necessary for these tests, and the groups comprised service personnel, research workers, students and members of W.E.A. classes. Both men and women took part, and it seems likely that the groups were, on the whole, above average in intelligence. It is probable that less intelligent subjects would not have understood the tasks.

One group was carefully selected, however, for ninety-three teachers were asked to do test E. This was because they were thought to be 'expert' in the use of such qualitative terms, which are often used for school marking. The data provided by this group were examined separately. The group consisted of both men and women teachers from five primary, three secondary modern and two grammar schools.

RESULTS

These will be discussed in three sections, section 1 dealing with the results obtained from test A, section 2 with the results from tests B, C and D, and section 3 dealing with the results from test E.

Section 1

Only ten subjects were asked the questions in test A, and this test did little more than demonstrate how difficult subjects found such a problem, particularly when the situation was left completely ambiguous.

The answers varied enormously, 'Very far' being given as 100 miles by one person, and 1000 light-years by another, and 'Very near' being given as 200 yards by one person, and the diameter of a molecule by another.

There was one other interesting feature, in that the subjects often gave responses in an unexpected order, thus the distance for 'Rather far' would sometimes be greater than the distance for 'Far', and sometimes less.

Section 2

In tests B, C and D the situations were more carefully defined, but no more carefully than they would be in everyday conversation. Seventeen subjects who did these tests were presented with rating scales which were 'inconsistent' in that for one of the three tests the 'Rather' term on the scale was written in next to 'Very', instead of next to the term 'neither x nor y '.

From the results, it was obvious that most people of this group thought that 'Rather' should come next to 'neither x nor y ', although only four out of seventeen subjects were careful enough to give this order consistently for all tests. Evidently, then, the word 'Rather' is an ambiguous one, and its interpretation will depend largely on the position which is given to it in the rating scale presented to the subjects.

Sixty-eight subjects were given tests B, C and D with the descriptive phrases in the more generally approved order, that is, with the 'Rather's' next to 'neither x nor y '.

A histogram for the results of test D, for these sixty-eight subjects, is shown in Fig. 1, which the number of pounds, grouped logarithmically, are plotted against the frequency responses for each of the descriptive terms. It can be seen that there is very considerable disagreement about the value for each of the terms.

There is one consideration to be taken into account, however, in that test D, in common with tests A, B and C, asks subjects to define a scale which is unlimited at one end, but limited at the other. Thus, for example, all subjects would agree that a sum of £1000 per

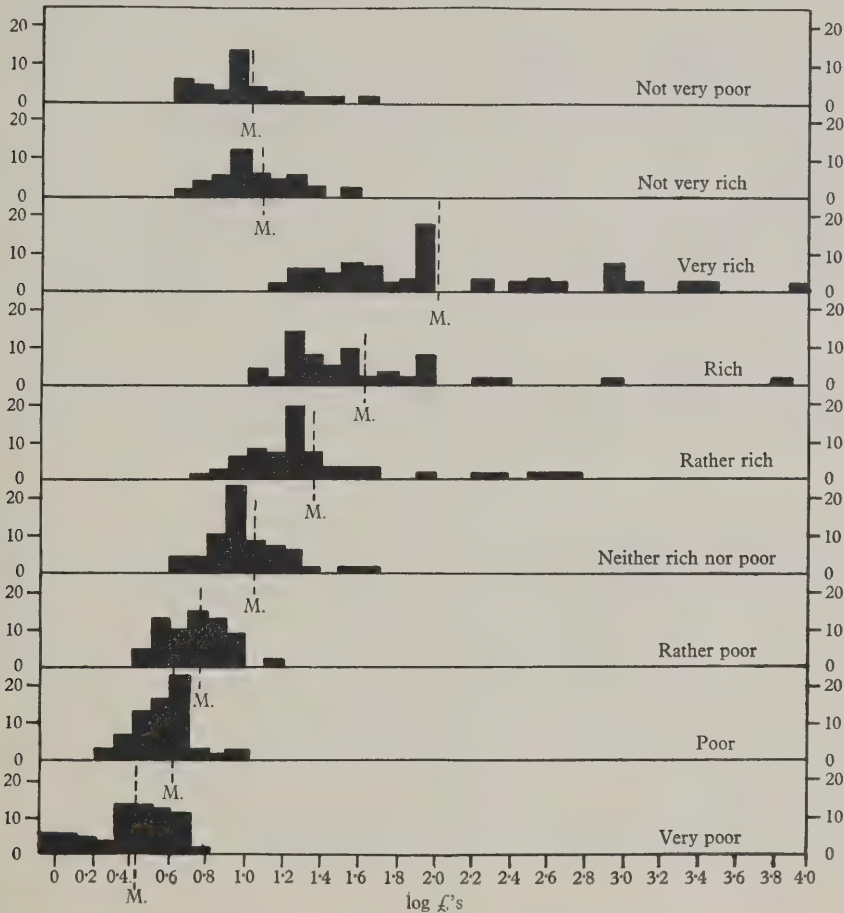


Fig. 1. Frequency distribution of responses for test D.

week would make a single man 'Very rich', although many subjects gave less than this sum. In addition, there is naturally more scope for disagreement at the 'unlimited' end of the scale than the 'limited' end.

Table 1 shows the range of the responses for test A and the logarithmic means and standard deviations for tests B, C and D as done by the sixty-eight subjects. The extreme values which were given for these latter tests are also shown.

Plotting the data logarithmically was a reasonably successful procedure with the results given for tests B, C and D, but most successful with test D, where subjects tended to give values first of all in pounds, then in tens, in hundreds, and finally in thousands

Table 1. *The variations in the objectivity measurements given for qualitative values*

Test A	Descriptive term	Very near	Near	Range of values given				Far	Very far
				Rather near	Neither near nor far	Rather far	Far		
Test B	Smallest value	4×10^{-9}	1 yd.	$\frac{1}{4}$ mile	150 yd.	6 miles	2 miles	100 miles	—
	Greatest value	200 yd.	2 miles	20 miles	63 miles	To the moon	To the sun	1,000 light-years	—
Test C	Descriptive term	Very near	Near	Rather near	Neither near nor far	Rather far	Far	Very far	Not very far
	Smallest value	2 yd.	15 yd.	20 yd.	300 yd.	440 yd.	$\frac{3}{4}$ mile	3 miles	300 yd.
	Greatest value	3 miles	5 miles	7 miles	10 miles	20 miles	50 miles	100 miles	12 miles
Test D	Descriptive term	Very short	Short	Rather short	Neither short nor long	Rather long	Long	Very long	Not very long
	Smallest value	1 sec.	30 sec.	1 min.	3 min.	5 min.	10 min.	15 min.	$2\frac{3}{4}$ min.
	Greatest value	5 min.	10 min.	20 min.	30 min.	45 min.	2 hr.	3 hr.	30 min.
Test E	Descriptive term	Very poor	Poor	Rather poor	Neither rich nor poor	Rather rich	Rich	Very rich	Not very rich
	Smallest value	8s.	£2	£3	£5	£7	£12	£16	£4. 10s.
	Greatest value	£5. 6s. 7d.	£10	£15	£50	£500	£4,000	£20,000	£40

Means and standard deviations for values grouped logarithmically

Test B	Descriptive term	Very near	Near	Rather near	Neither near nor far	Rather far	Far	Very far	Not very far	Not very near
Test C	Log mean (miles)	1.11	1.59	1.97	0.32	0.64	0.87	1.10	0.16	0.30
	Mean	228 yd.	677 yd.	1640 yd.	2 miles 156 yd.	4 miles 625 yd.	7 miles 737 yd.	12 miles 985 yd.	1 mile 755 yd.	1 mile 1735 yd.
Test D	Log s.d.	0.67	0.56	0.48	0.40	0.35	0.30	0.38	0.23	0.39
	Descriptive term	Very short	Short	Rather short	Neither short nor long	Rather long	Long	Very long	Not very long	Not very short
Test E	Log mean (min.)	1.89	2.37	2.71	2.88	3.15	3.38	3.61	2.92	3.03
	Mean	47 sec.	2 min. 21 sec.	5 min. 8 sec.	7 min. 35 sec.	14 min. 0 sec.	23 min. 59 sec.	40 min. 45 sec.	8 min. 55 sec.	10 min. 43 sec.
Test F	Log s.d.	0.52	0.50	0.26	0.24	0.23	0.23	0.30	0.20	0.35
	Descriptive term	Very poor	Poor	Rather poor	Neither rich nor poor	Rather rich	Rich	Very rich	Not very rich	Not very poor
Test G	Log mean (£'s)	0.42	0.61	0.78	1.04	1.38	1.61	2.01	1.07	1.01
	Mean	£2. 12s. 4d.	£4. 16s. 7d.	£6. 5s. 2d.	£10. 19s. 1d.	£23. 15s. 5d	£40. 14s. 9d.	£101. 8s.	£11. 13s. 5d.	£10. 4s. 1d.
Test H	Log s.d.	0.20	0.14	0.18	0.19	0.41	0.41	0.61	0.19	0.23
	Descriptive term	Very poor	Poor	Rather poor	Neither rich nor poor	Rather rich	Rich	Very rich	Not very rich	Not very poor

Table 2. *Means and standard deviations for 'abstract' scale (test E)*

Test E	Descriptive term	Very bad	Bad	Rather bad	Average	Rather good	Good	Very good	Not very good	Not very bad
		Mean in inches	s.d. in inches	Mean in inches	s.d. in inches	Mean in inches	s.d. in inches	Mean in inches	s.d. in inches	Mean in inches
Test I	Mean in inches	6.3	5.5	4.9	3.4	2.3	1.7	0.8	3.4	3.9
	s.d. in inches	0.32	0.44	0.70	0.29	0.47	0.67	0.27	0.57	0.58

Number of subjects 78

of pounds. With test B, where the scale started off in yards, through fractions of miles, miles and tens of miles, this conversion was not so successful.

Evidently then, the fact that certain expressions come more readily to mind than others limits the value of data obtained from such an experiment. As an example, subjects all describe the earnings of a 'Rich' single man in multiples of £5.

Even despite this consideration, however, it remains true that the values given for the contiguous qualities are logarithmically related to one another, at least approximately, so that this might indicate that the Weber law, or something like it, can be applied to real-life mental schema.

As Wolters (1916) has suggested, the negative concepts proved to be interesting in these tests. Only forty out of the sixty-eight subjects inserted the negative phrases at all in test D, and of these eleven considered that to be 'Not very poor' was about the same as being 'Not very rich', twenty-three considered that to be 'Not very poor' implied that one was poorer than 'Not very rich' and six held the contrary opinion.

'Not very rich' was held to mean anything between 'Very rich' and 'Rather poor', with thirty-three subjects placing this between 'Rather rich' and 'Rather poor'; 'Not very poor' was held to mean anything between 'Rich' and 'Poor', but thirty-seven placed it between 'Rather rich' and 'Rather poor'.

When the results for different scales were compared for different individual subjects, however, it was found that the negative judgements were not consistently placed, either in their relation to one another, or in their relation to the positive judgments.

Section 3

Excluding the teachers, 118 subjects were given these particular phrases for test E; of these some 73% used the order V.G., G., R.G., Av., R.B., B., V.B. Histograms for the results are shown in Fig. 2; in which the position of the intersection on the 7 in. line is plotted against the frequency of occurrence.

The means and standard deviations for these data have been calculated since the distributions are in most cases approximately normal about a central point. These are given in Table 2. In some cases these may be misleading, however, since for certain qualities the distributions appear to be skewed.

No doubt some of the apparent disagreement results from subjects inserting their points inaccurately, but even with the term 'Average', for which there is more agreement than for any other, it is evident that some subjects cannot have intended that it should come at the mid-way position.

As before, there is less agreement about the placing of the negative terms, than for the positive ones.

Some 14% of these subjects gave the order V.G., R.G., G., Av., B., R.B., V.B., some 3% omitted certain of the terms since they did not feel sure of them, and some 10% were suggestible enough, or careless enough, to use the order given in the instructions, i.e. V.G., R.G., G., Av., R.B., B., V.B., an order that had been chosen so as not to influence the subjects' placing of 'Rather'.

Plotting the frequencies in a histogram form obscures the individual differences in the results. A large number of the individual results were therefore plotted out, as in Fig. 3; but it was not possible to find sufficient common features to be able to group them satisfactorily in any way. Some of the curves which resulted were linear and some were

approximately sigmoid, or inverted sigmoid, in shape; but for the majority it was only possible to say that the 'Bad' side was a mirror image of the 'Good' side. Some were completely irregular, as for example scale (d) in the figure.

There were no apparent differences in the subjective scales when the terms 'Fairly' or 'Quite' were substituted for the word 'Rather', except that there was more agreement that the former terms should come next to 'Average' than there was for the latter. The completed scales for which 'Moderate' and 'Fairly' were included were also very similar to those given for the other terms.

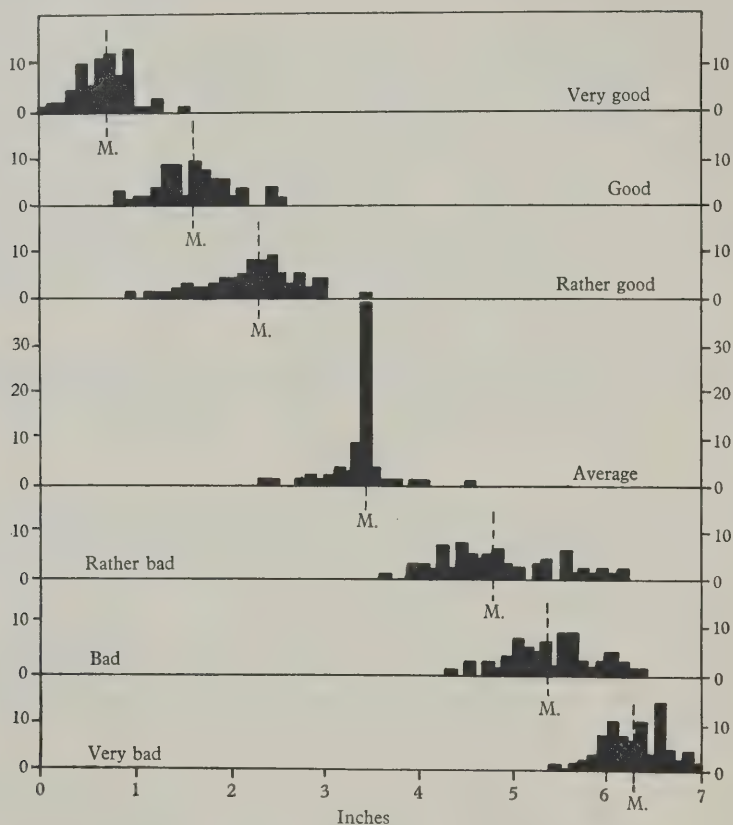


Fig. 2. Frequency distributions of responses for test E.

It is interesting, however, that although there was, in a sense, agreement between the groups, yet the resulting scales, when averages were plotted, were not linear ones.

Some subjects criticized these tests as being too abstract, but in fact they did seem to have some meaning, since it would have been possible to predict with some success the relation between the values given in tests B, C and D for different individuals, from their scales given by this test. Furthermore, it was possible to make a much smoother curve relating subjective judgements on dairy products to an equivalent objective measurement, by using a non-linear scale derived from tests of this type, than by treating the subjective values as though they were equally spaced from one another. In addition, the smoother curve was more like that which would have been expected if the Weber law was in operation (Prentice, 1953).

By way of interest, a large number of teachers were asked to do test E, since school marking is often done in such qualitative terms. Forty-three teachers were given the scale in which 'Rather good' and 'Rather bad' were included, but since they objected to the use of these terms, a further fifty teachers were asked to do the test with the term 'Fairly' substituted for 'Rather'. A few teachers even objected to this, and said they would use 'Fair' rather than 'Fairly bad', while nearly all objected to the use of the negative terms.

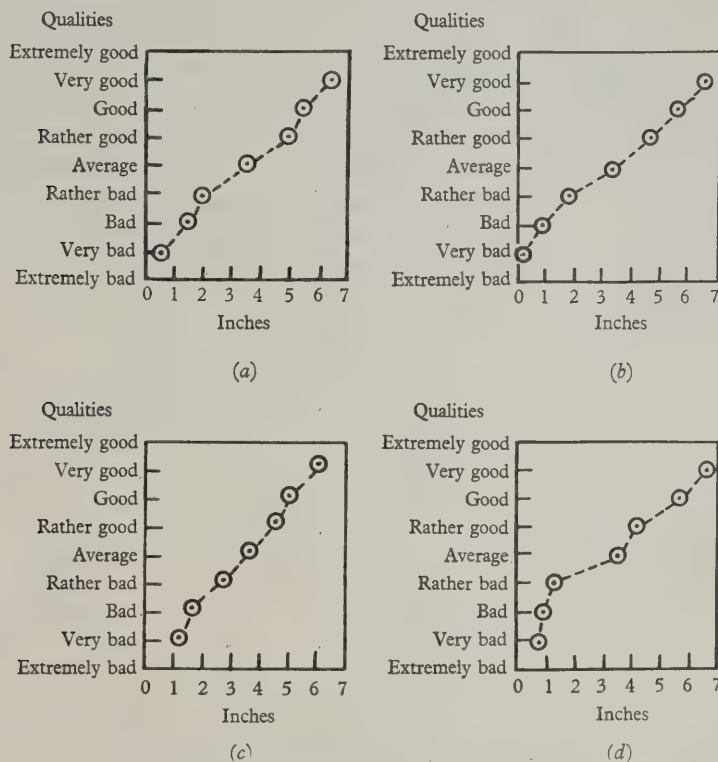


Fig. 3. Examples of subjective scales given for test E.

The scales which resulted were, on the average, only a little more linear than those given by other groups, and there was no more agreement about the positions on the line or the terms, between teachers in different schools, or even in the same schools, than there had been with the other groups of subjects. In Fig. 4, three scales given by teachers in the same school are shown, and it is easy to see that a child would be less likely to get a good mark from teacher C than from teacher A.

It was interesting to see subjects' reactions to these problems; thus some took the task very seriously, and made several attempts before being satisfied; some at once thought of reasons why it could not be done, or why the results would be meaningless; and a small number refused to, or were quite unable to do it. The latter were mostly people who take themselves rather seriously.

A good many points of interest were raised in comments. Thus three subjects thought that 'Bad' should have been on the left, and 'Good' on the right; one subject thought that the descriptions were not all in the same dimension; some defined the qualities by

marking in two points, the quality in question applying over this distance; and very many objections were raised to the negative terms.

DISCUSSION

Evidently the results of these tests show that such quantitative terms for describing qualities are very ambiguous; yet for certain purposes, as for example with rating scales, it may be better to use such terms, which everyday people can understand, than to ask judges to attempt the more abstract task of judging by a points system.

A method such as that used in test E allows a somewhat less arbitrary interpretation of such quality descriptions and, as previously stated, it has been found to be useful in relating subjective assessments to objective measurements. Furthermore, it seems that by using the somewhat ambiguous terms 'Good' and 'Bad' one can build up a scale which is as limited in one direction as it is in the other; while it is also possible to obtain a relation between the quantitative terms which is sufficiently independent of the quality (in this case 'Good' or 'Bad') for it to be possible to transfer this relation and apply it successfully to assessments for another quality. Thus this was possible in the examples

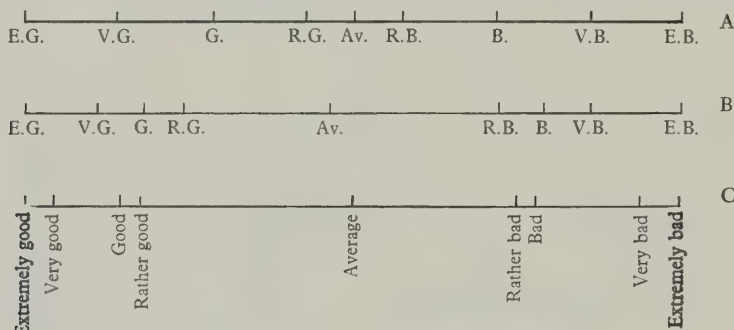


Fig. 4. Scales given by three teachers.

given here, for the subjective tests on food qualities, and for the qualities represented in tests B, C and D.

It can be seen from the histograms in Figs. 1 and 2 that if *complete* agreement is required between subjects, nothing but a three-point scale using the two 'Very' terms, and 'Average' could give this. But a technique such as that for test E does at least make it possible to interpret the meaning of descriptive terms given by individuals, and to weight them so that the use of all seven points on a seven-point scale is at least a little less arbitrary.

One can, therefore, transfer the results of a technique such as method E to other subjective assessments, where more satisfactory verbal descriptions of the different qualities may be impossible, although such a transfer does involve making certain assumptions.

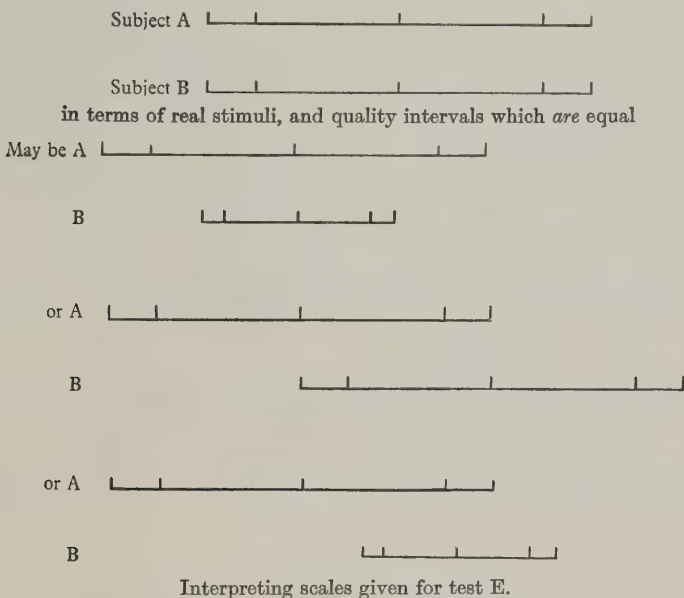
Thus, for example, it can only show the relative positions of the quality steps, one to another. There is no way of knowing from this whether subjects would agree about 'equal quality steps'; whether all judges would agree that a certain item is 'Very good', even if they agree to place 'Very good' in the same relative position to 'Extremely good'; or how long the scales are in relation to one another, even if the subjects agreed about 'equal quality steps'. Fig. 5 shows the sort of difficulty that is meant.

It might be possible to determine some of the answers to such problems by combining data obtained from a method such as test E with an examination of ratings on identical samples by different judges, though this would not be easy to do.

However, as previously stated, by choosing two values, such as 'Extremely good' and 'Extremely bad', and assuming that judges agree about these, one can at least make interpretation of subjective judgements one stage less arbitrary than by assuming that the qualities are equally spaced one from another. If nothing else, it reduces the number of assumptions and, apparently, one can improve the interpretation of subjective assessments by such methods.

It is clear, of course, that although such methods may make it possible to interpret the popular meaning of quantitative descriptive terms more successfully, it will still be necessary to select the other criteria, to which these terms will be applied, as carefully as possible (Hopkinson, 1950).

Fig. 5. Identical scales between arbitrary points, obtained by the method of test E



The results emphasize, however, that assessments made in terms of quantitative descriptions such as these should not be treated as if they were linearly related to one another. Furthermore, while it is true that subjects are to some extent suggestible, and will be affected by the order in which the descriptive qualities to be used in the rating scale are written out, since the only feature common to the results given for test E was that subjects inserted the phrases at positions which on the 'Bad' side, were a mirror image of those on the 'Good' side, it would seem better to construct rating scales so that the descriptive terms on each side of 'Average' also mirror each other.

SUMMARY

1. Some simple experiments were carried out to see how far people agreed in their use of adjectives qualified by quantitative adverbs, both for measurable items of everyday experience, and by a more abstract method.

2. The results suggest that if complete agreement, with no overlapping, is required, only three points on a 7-point rating scale of quality descriptions should be used.

3. A method is suggested by which the procedure of giving numerical values to these quality descriptions can be made one stage less arbitrary, and some experimental verification for such a method is quoted.

The author wishes to express his gratitude to his many subjects for their co-operation and, in particular, to the headmasters of the various schools for giving him permission to question their staffs. In addition, he wishes to acknowledge his debt to Dr G. W. Scott-Blair, Prof. R. C. Oldfield, Mr J. H. Prentice and Miss S. D. Carrinci for helpful discussions; to the latter for invaluable help in the analysis of the data; and to the Agricultural Research Council for a grant.

REFERENCES

- HOLLINGWORTH, H. L. (1910). The central tendency of judgment. *J. Phil. Psychol. Sci. Meth.* vii, 461.
 HOPKINSON, R. G. (1950). The multiple criterion technique of subject appraisal. *Quart. J. Exp. Psychol.* ii, 124.
 PRENTICE, J. H. (1953). The effect of storing and blending on the firmness and spreadability of butter. *Proc. XIII Dairy Congr.* ii, 723.
 SHEPPARD, D. (1953). Subjective assessments of firmness. The use of a rating scale. *Quart. J. Exp. Psychol.* v, 1.
 TRESSELT, M. E. (1948). The effect of the experiences of contrasted groups upon the formation of a new scale of judgment. *J. Soc. Psychol.* xxvii, 209.
 WOLTERS, A. W. P. (1916). The process of negation. *Brit. J. Psychol.* viii, 183.

(Manuscript received 20 August 1953)

VERBAL EXTRAPOLATION AND INTERPOLATION

By E. C. POULTON

Medical Research Council Applied Psychology Research Unit, Cambridge

I. *Introduction* (p. 51). II. *Method* (pp. 51–52). III. *Scoring* (pp. 52–53). IV. *Results* (pp. 53–55). V. *Discussion* (pp. 56–57). VI. *Summary* (p. 57). *References* (p. 57).

I. INTRODUCTION

Experimental and statistical investigations of verbal redundancy (e.g. Miller, 1951; Newman & Gerstman, 1952) have used the method of extrapolation or forward-association. They have answered the question: 'How much new information is *added* by the next word or letter in a message?' This has been deduced from the frequency with which the component follows the previous one, two, three, ..., components. However, the method of interpolation or forward- and backward-association can also be used, and may in certain circumstances be more relevant, as it answers the question: 'How much new information is *given* by a particular component of a message?' This depends upon the frequency with which the component is found in its position in the particular type of passage, the component having to be related both to what has preceded it and to what follows it.

There are situations in everyday life permitting verbal extrapolation but not interpolation, as when a telephone conversation is cut off in the middle. But there are other situations permitting either, as when a sudden loud noise blots out the middle part of a message. Which of the two occurs in these circumstances can only be decided by experiment.

The original aim of the present experiments was to determine whether short gaps in verbal material are filled simply by extrapolation, or by interpolation. A subsequent aim was to discover whether by a suitable choice of the words immediately succeeding a short gap, interpolation could be made to give a wider range of answers than extrapolation.

II. METHOD

Extrapolation experiment. Of the thirty subjects, three were tested at a time, seated at tables screened from each other. A light short story of about 2000 words was played to them from a tape recorder. At 1 min. intervals the story was stopped. The thirty-one stopping places were selected on a time basis, independently of the content of the story. They were each marked by a sharp tap, and were followed by a 15 sec. pause. During the pause subjects had to write down the next two or three words of the story. They were instructed to answer as quickly as possible. Any answer was as good as any other, provided it fitted the context. Delays of more than 5 sec. in starting to write were noted by the experimenter, using a stop-watch. There was a short practice before the experiment, to ensure that subjects had understood the instructions.

Of the thirty-one stopping places, ten were at the start of a sentence, six of these after a link with the previous sentence such as 'and', the other four after the subject of the sentence; thirteen were after the subject and verb; and eight were after the subject, verb and object. The words immediately before a stopping place will be called the *precontext*.

Story-postcontext experiment. The procedure was similar for a new but comparable group of thirty subjects, except that they had to interpolate. The same thirty-one stopping places were each marked by a loud buzz lasting 1-2 sec. This was followed by a few more words taken from the story, the *postcontext*, before the 15 sec. pause. During the pause subjects had to write down the words omitted during the buzz.

The size of the actual gap left in the story ranged from one to eight words, the average being 4.1; but by reducing redundancy it could be filled fittingly by an average of 1.8 words. The start of the postcontext was chosen so that its meaning and grammatical structure affected the filling of the gap as little as possible. Generally it was a new sentence or a clause. In only two cases did the postcontext start after the subject of the sentence, so that this necessarily became part of the gap to be filled. The postcontexts ranged from 3 to 17 words, the average being 8.5.

Changed-postcontext experiment. In a second interpolation experiment nineteen of the postcontexts were altered in order to change the nature of the answers given by the group of thirty new but comparable subjects. A representative sample of twelve postcontexts was left unchanged. The mean postcontext length was reduced slightly to 7.9 words, the range being 2 to 16. The gaps could be filled by an average of 1.9 words.

Subjects. The ninety subjects were all naval ratings from Fleet Air Arm stations, aged between 17 and 31 years. Most had left school at 14 or 15 and had become apprentices or taken miscellaneous jobs before being called up for military service. Each performed the experiment once only.

III. SCORING

Commonest meaning changes. The number of extrapolation subjects giving the commonest meaning of their experiment at each stopping place, was compared with the number of subjects in the two interpolation experiments giving this meaning. Similarly, the number of story-postcontext subjects giving their commonest meaning (which might or might not be the same as for the extrapolation experiment) at each stop, was compared with the number from the changed-postcontext experiment. Classifying answers by meaning sometimes seemed quite arbitrary, especially at stops giving answers with many different meanings. In these cases meaning categories were widened until the commonest meaning of the first experiment was given by at least seven out of the thirty subjects (25%), in order to ensure that the classification was reasonably representative of the answers given.

Changes in the incidence of commonest meanings in the comparison groups were predicted for one of three reasons: (a) the commonest meaning and the new postcontext meaning were in conflict (excessive redundancy, or contradiction); (b) the new postcontext demanded another meaning, the commonest meaning being unnecessary but possible in addition; and (c) the new postcontext demanded this meaning. In the latter case an increase was predicted, in the two former a decrease. These prediction criteria were formulated while comparing the extrapolation and story-postcontext answers. Only the other two comparisons are therefore shown in the results section below.

Changes in first word of answer. The same comparisons were made for commonest first words of answers as for commonest meanings. Changes in incidence were predicted using three similar criteria. Only the verbal structure of the postcontext was considered in this case, not its meaning. Where two first words were equally common in the first experiment, half weight was given to each. Nearly identical words, e.g. singulars and plurals, *a* and

in, round and *around*, were grouped together. In the extrapolation-interpolation comparisons one stopping place had to be scrapped, as the precontext of the extrapolation experiment ended in the middle of a word, and was therefore changed for the two interpolation experiments.

Four-fifths of the commonest first-word predictions occurred at the same stops, and were in the same directions, as the commonest meaning predictions. The example which follows (stopping place no. 30) was chosen to show as many of the points mentioned above as possible:

Extrapolation experiment: 'but he wanted more (information), and he wheeled.....'. Observed commonest meaning: on some member of the crew. Observed commonest first word: around (or round).

Story-postcontext experiment: 'and he wheeled.....for the required illumination'. Observed commonest meaning: around in general, not on a member of the crew. Observed commonest first word: around (or round).

Changed-postcontext experiment: 'and he wheeled.....of the cook for the required illumination'. Predicted changes in extrapolation experiment commonest meaning: decrease, because 'cook' in postcontext makes a member of the crew redundant. Predicted changes in extrapolation experiment commonest first word: decrease, because 'around' is unnecessary, while the new postcontext requires something like 'in the direction'.

Predicted changes in story-postcontext experiment commonest meaning: nil. Predicted changes in story-postcontext experiment commonest first word: decrease.

The change in the total number of different first words used at each stopping place was assumed to vary inversely with the change in the incidence of the commonest first word, and predictions were made accordingly. The contexts and predictions at the different stops are given in detail in Poulton (1953), together with a full analysis of the answers at one stop.

Indices of difficulty in answering. Omissions include failures to answer, and cases where subjects did not start to write until the story had recommenced after the 15 sec. pause. Delays are restricted to cases where subjects started to write between 5 and 15 sec. after the beginning of the pause. The number of answers which conflicted either in grammar or meaning with the precontext, or the postcontext when present, were also noted.

Calculations. Whenever possible the significance of the difference of each pair of means was assessed separately against the variability of the subjects for all stopping places in the group concerned, as well as against the variability of the stopping place differences for all the subjects combined. However, the significance levels of the latter comparisons have always been taken, as they were the less significant. Levels beyond $P=0.001$ will be called 'highly significant', beyond $P=0.01$ 'significant', and beyond $P=0.05$ 'just significant'. Both tails of the assumed normal distribution have always been used.

IV. RESULTS

Extrapolation and changed-postcontext experiments

(1) The incidence of the commonest meanings of the extrapolation experiment is shown in Table 1 for this experiment and for the changed-postcontext experiment. Each stop has been allocated to one of three groups, depending upon whether its changed post-

context should have altered the incidence or not, and upon the predicted direction of the alteration, using the standard criteria given above. The differences between the two experiments and their p values are also shown.

(2) The incidence of the commonest first words of the extrapolation experiment is shown in Table 2 for the two experiments. Again each stop has been allocated to one of three prediction groups, but as compared with Table 1, a few of the stops have changed groups.

Table 1. *Extrapolation and changed-postcontext experiments:
commonest extrapolation meanings*

	Changed postcontext predicted to produce		
	No change (17 stops)	Decrease (12 stops)	Increase (2 stops)
Extrapolation (%)	51	52	40
Changed-postcontext (%)	44	16	75
Difference (%)	7*	36	-35
P	0.05	0.001	0.05

* No change—decrease, $p=0.001$; no change—increase, $p=0.001$.

Table 2. *Extrapolation and changed-postcontext experiments:
commonest extrapolation first words*

	Changed postcontext predicted to produce		
	No change (18 stops)	Decrease (11 stops)	Increase (1 stop)
Extrapolation (%)	45	46	33
Changed-postcontext (%)	39	18	50
Difference (%)	6	28*	-17
P	0.05	0.01	—

* No change—decrease, $p=0.01$.

Table 3. *Extrapolation and changed-postcontext experiments:
different first words*

	Changed postcontext predicted to produce		
	No change (18 stops)	Increase (11 stops)	Decrease (1 stop)
Extrapolation (%)	36	31	53
Changed-postcontext (%)	37	45	30
Difference (%)	-1	-14*	23
P	>0.05	0.05	—

* No change—increase, $p=0.01$.

(3) The total number of different first words used in the extrapolation and changed-postcontext experiments is shown in Table 3. The prediction groups are identical with those of Table 2, but the increase and decrease predictions have been reversed.

(4) The changed-postcontext experiment gave significantly more delays in answering than the extrapolation experiment (14% as compared with 5%). It yielded significantly more answers which did not fit the precontext in grammar and/or meaning (4% as compared with 1%). In both cases the stops at which no change in answer was predicted showed just significant differences.

Story-postcontext and changed-postcontext experiments

(5) The incidence of the commonest meanings and commonest first words of the story-postcontext experiment is shown in Tables 4 and 5 for this experiment and for the

changed-postcontext experiment. The stops with postcontexts common to both experiments are shown separately, and the remaining stops have been allocated to one of two groups, depending upon whether the changed postcontexts should have decreased the incidence or not.

Table 4. *Story-postcontext and changed-postcontext experiments:
commonest story-postcontext meanings*

	Same postcontext (12 stops)	Changed postcontext predicted to produce	
		No change (9 stops)	Decrease (10 stops)
Story-postcontext (%)	56	46	58
Changed-postcontext (%)	45	63	17
Difference (%)	11*	-17	41
<i>P</i>	0.01	0.05	0.001

* Same postcontext—no change predicted, $p=0.001$; same postcontext—decrease predicted, $p=0.001$.

Table 5. *Story-postcontext and changed-postcontext experiments:
commonest story-postcontext first words*

	Same postcontext (12 stops)	Changed postcontext predicted to produce	
		No change (12 stops)	Decrease (7 stops)
Story-postcontext (%)	44	42	55
Changed-postcontext (%)	37	38	26
Difference (%)	7	4	29*
<i>P</i>	>0.05	>0.05	0.05

* Same postcontext—decrease predicted, $p=0.05$.

Table 6. *Story-postcontext and changed-postcontext experiments:
different first words*

	Same postcontext (12 stops)	Changed postcontext predicted to produce	
		No change (12 stops)	Increase (7 stops)
Story-postcontext (%)	37	37	27
Changed-postcontext (%)	38	38	44
Difference (%)	-1	-1	-17*
<i>P</i>	>0.05	>0.05	0.01

* Same postcontext—increase predicted, $p=0.05$.

(6) The total number of different first words used in the story-postcontext and changed-postcontext experiments is shown in Table 6. The prediction groups are identical with those of Table 5, the predicted decrease having become a predicted increase.

(7) The changed-postcontext experiment gave significantly more delays than the story-postcontext experiment (14% as compared with 7%). It also yielded just significantly more answers not fitting the postcontext (17% as compared with 9%). Answers not fitting the precontext remained unchanged.

(8) Omissions showed no significant differences at all, always lying in the region of from 0 to 4%.

V. DISCUSSION

The influence of postcontext. In Tables 1 and 2 the extrapolation and changed-postcontext experiments have been compared on the commonest meanings and commonest first words of the extrapolation experiment. This experiment would be expected to yield a slightly higher incidence than the other by chance alone (just as a bridge player's longest suit in one hand tends to be longer than this same suit in his next hand). This expectation is borne out by the scores at the stops where no change was predicted, for in both tables the extrapolation experiment shows a just significant advantage. It distorts direct comparisons between the two experiments at the stops at which changes were predicted. However, the differences between the experiments at the predicted no-change and predicted decrease and increase stops can be compared, because this trend should affect all three groups of stops about equally. In both Tables 1 and 2 the relationships between these differences are as predicted, the p values being given as footnotes.

A direct comparison can be made between the extrapolation and changed-postcontext experiments in the total number of different first words used. Table 3 shows that there was no difference at the stops at which no change was predicted, but differences as predicted at the remaining stops.

Tables 1-3 thus show that changing a sentence continuation task from extrapolation to interpolation can alter the answers given, provided the postcontexts are suitably chosen as they were at the prediction stops of the changed-postcontext experiment. Result 4 shows that even when the postcontext did not alter the answers appreciably, it tended to add to the difficulty of answering, thus indicating that some attention was being paid to it. Clearly, when a short section of a verbal message is lost in transmission, the gap will be filled by reference to what follows it as well as to what precedes it. The forward-association method of study (Miller, 1951, pp. 187-8; Newman & Gerstman, 1952) would thus be an over-simplification of this situation.

The influence of the postcontext is also shown in Tables 4-6, in which the original story postcontexts are compared with the changed postcontexts. In tables 4 and 5 it is the differences between the story-postcontext and changed-postcontext experiments which should be compared (cf. discussion of Tables 1 and 2 above). In both tables the relationship between the same postcontext and predicted decrease groups of stops is as predicted (see p values in table footnotes). In Table 6 a direct comparison can be made between the two experiments, and again the difference is as predicted.

The high incidence of the commonest story-postcontext meanings found in the changed-postcontext experiment at the stops at which no change was predicted (Table 4) requires comment. At seven of the stops the changed postcontext appeared to suggest this meaning more strongly than the corresponding original story postcontext. However, none of the differences in postcontext was sufficiently definite for prediction to be possible using the standard criteria given above.

Context and range of answer. Many of the changed postcontexts were designed not to fit the obvious meanings and/or first words suggested by the precontexts. At these stops the changed-postcontext experiment yielded less uniformity in answer than the extrapolation experiment. Table 3 shows the resultant increase in the number of different first words used (predicted increase stops). The decrease in the commonest extrapolation meanings (Table 1 predicted decrease stops) suggests that there was probably a corresponding

increase in the range of meanings, although no satisfactory measure of this could be devised. This reduced uniformity was not due simply to an increase in answers not fitting the precontext, for the latter was small and was spread more or less evenly over all the tops (Result 4).

In these cases the additional constraint imposed by the postcontext increased the range of the answers. Thus by *adding* to a context it is possible to *reduce* its effectiveness in signposting the answers. This emphasis on reduced uniformity of answer, rather than on increased uniformity in the design of the changed postcontexts, must account for the greater difficulty of this experiment as compared with the story-postcontext experiment (Result 7).

VI. SUMMARY

A comparison was made between continuing sentences (extrapolation) and filling gaps of unstated size in sentences (interpolation), when both operations occurred at the same randomly selected places in a story. It was found that the operations did not necessarily give different answers, although interpolation was more difficult than extrapolation. However, by a suitable choice of postcontexts, changes could be produced in the meanings and first words of the answers. The addition of a postcontext could actually increase the range of the answers given.

The influence of Prof. Sir Frederic Bartlett, F.R.S., is clearly shown in this paper. Statistical advice was received from Mr N. L. Webb. The subjects were supplied by the Royal Navy. Financial support from the Medical Research Council is also gratefully acknowledged.

REFERENCES

- HILLER, G. A. (1951). *Language and Communication*. New York: McGraw-Hill.
 LEWIS, E. B. & GERSTMAN, L. J. (1952). A new method for analysing printed English. *J. Exp. Psychol.* XLIV, 114-25.
 POULTON, E. C. (1953). Verbal extrapolation and interpolation. *Cambridge Appl. Psychol. Unit, Rep.* no. 172.

(*Manuscript received 5 September 1953*)

A PRELIMINARY INVESTIGATION OF THE EFFECTS OF CONTINUOUS STIMULATION ON THE PERCEPTION OF 'APPARENT MOTION'

By M. B. SHAPIRO

Department of Psychology, Institute of Psychiatry, Maudsley Hospital

I. *Introduction* (p. 58). II. *Explanatory hypotheses and findings* (pp. 58-60). III. *Deductions from the hypotheses* (p. 60). IV. *The experiment* (pp. 60-62): (i) *Purpose* (pp. 60-61); (ii) *Apparatus* (p. 61); (iii) *Sample* (p. 61); (iv) *Procedure* (pp. 61-62). V. *Results* (pp. 62-64). VI. *Discussion* (p. 64). VII. *Alternative explanations of the findings* (pp. 64-67): (i) *Displacement effect* (pp. 64-65); (ii) *A lowered intensity of apparent stimulation* (pp. 65-67). VIII. *Summary* (p. 67). *References* (p. 67).

I. INTRODUCTION

The experiment reported in this paper is part of an investigation into the psychological effects of brain damage. The particular observation which was the starting-point for this experiment was one reported by Werner & Thuma (1942) who found that, under certain conditions, brain-damaged subjects failed to see apparent motion.¹ The conditions consisted of presenting the subjects, by means of a modified Dodge tachistoscope, with successive exposures of two vertical black lines. At a certain speed of succession all but one of twenty normal subjects saw movement. Only two of twenty brain-damaged subjects saw movement.

Both groups were of low I.Q., the average of the brain-damaged group being 69, and of the non-brain-damaged group being 72. It is stated that the brain-damaged group did not suffer from any gross motor defects or visual agnosias, although specific diagnoses of the type of brain damage suffered by these subjects are not stated.

A related finding is that of Bender & Teuber (1949), who found that patients with occipital-parietal lesions failed to perceive apparent motion when certain affected parts of the visual field were appropriately stimulated.

II. EXPLANATORY HYPOTHESES AND FINDINGS

Two hypotheses were formulated in explanation of this finding:

(i) The first was in essence the same as Wertheimer's explanation of the occurrence of apparent motion in normals. According to this hypothesis visual stimulation in any relevant part of the brain is followed by the irradiation of excitatory effects from the site of stimulation. If such stimulation is followed by stimulation in another part of the brain, sufficiently soon after and appropriately placed, the irradiation effect of the second site of stimulation is presumed to combine with the irradiation effect of the first in some way so as to produce the visual perception.

It should be noted that it is not the full and detailed Wertheimer hypothesis which is being formulated here. All that is being used is the general notion of the irradiation of excitatory effects and their combination to provide the basis of the perception of apparent motion.

¹ This finding was brought to my attention by Mrs M. Brenner.

Some experiments have produced results which appear to throw considerable doubt on his theory. Closer examination of these experiments shows, however, that they either are based on incorrect assumptions or deal only with secondary aspects of the Wertheimer hypothesis.

Two examples will be given in illustration. The first of these is an article by K. U. Smith (1940). He found that the same nystagmic head movements were produced in the guinea-pig by real motion and 'apparent motion'. These movements were not affected by the removal of the animal's cortices. This experiment, Smith believes, demonstrates the inadequacy of the Wertheimer theory. He was apparently unaware of the fact, discussed by Marquis (1935), that in the lower vertebrates the visual functions are carried out entirely by the mesencephalic centres, while in man the cortex has taken over all the visual functions with the exception of the pupillary reflex to light.

Our second example of a claimed disproof of the Wertheimer hypothesis is a paper by R. Smith (1948). Ten human subjects were given successive visual stimulation so that each stimulus fell upon the retina of its corresponding eye at the nasal end of the horizontal meridian. The visual projection areas of these two parts are completely separated from each other by the longitudinal fissure of the cerebrum, one lying on the right and the other on the left occipital lobe. The afferent fibres leading to these projection areas are also separate. Thus the two experimental stimuli were represented respectively at the extreme margin of the left visual cortex and at an opposing point on the right visual cortex. Seven of the ten subjects reported complete horizontal movement. Smith concludes from this that the general theory of corresponding interaction is thrown into considerable doubt. The first point to be noticed is that none of Smith's subjects reported the unequivocal perception of object movement; only 'pure phi' was definitely seen. 'Pure phi' is an experience of movement in which no moving object is seen. The second objection to Smith's conclusion is that a connexion between the two occipital poles may exist via the corpus callosum and possibly the lateral geniculate bodies. Any two parts of the brain which are interconnected may interact.

In view of such considerations the necessity still remains for testing the Wertheimer hypothesis in the generalized form described above.

(ii) The second hypothesis was that one of the effects of brain damage on psychological processes is the increase of inhibitory effects in the affected area. Inhibitory effects are here defined as the demonstrable decrease or disappearance of excitatory effects as the result of some process in the brain. The intensity and range of inhibition will vary, in as yet undefined ways, with the site and extent of the lesion. The grounds for this hypothesis were the work of Teuber & Bender (1949) and Pavlov (1927).

The findings of Teuber & Bender (1949) were obtained on patients who had occipito-parietal lesions. They report that extinction occurs in mildly impaired areas of the visual field, as soon as the other relatively intact areas are stimulated simultaneously. Thus a pattern may be seen on exposure to an amblyopic half of a field, provided this exposure takes place against a relatively homogeneous background. However, as soon as the second pattern is simultaneously exposed in the other (less impaired) half of the field, the first pattern becomes 'extinct'. 'The patient states, often with signs of surprise, that he does not see the first pattern any longer. . . . However, as soon as the pattern in the less impaired part of the field is removed, the one in the affected field becomes visible.'

Pavlov's findings were obtained from dogs which had had part of their cortices removed.

For a period after operation, previously established conditioned reflexes related to the site of the injury could not be evoked. Instead, a definite inhibitory effect in the form of sleep was produced.¹

What is common to both these findings is that, in descriptive terms, inhibitory effects are increased as a result of brain damage. We are not concerned here with the mechanisms underlying these findings about which there is some disagreement but only with the facts at a descriptive level.

III. DEDUCTIONS FROM THE HYPOTHESES

From these two hypotheses it would follow that one possible explanation of the failure of the brain-damaged patient to see apparent motion is that the irradiation effects postulated by the first hypothesis above are inhibited, thus removing the basis for the perception of apparent motion.

One deduction which follows from this explanation is that the introduction of an inhibitory condition in normal subjects should result in their failing to see apparent motion at speeds of succession which normally would be sufficient. No assumption is being made here about the nature of the inhibitory processes involved. We are dealing here only with effects. Such a condition should result at least in the need to shorten the time interval between the two stimulations in order to produce the perception of motion.

In order to test this deduction it was necessary to find a method of producing an inhibitory condition in the brains of normal subjects. It was decided to do this by the method of continuous visual stimulation. This was suggested by the work of Kohler & Wallach (1944), who report that continuous visual stimulation can result in making the stimulus phenomenologically less intense. Such stimulation also has inhibitory effects on the surrounding areas of the brain, the effect decreasing as the distance from the point of stimulation increases. These facts have been demonstrated by Bevan (1951). He investigated the changes in the intensity thresholds of the perception of a light stimulus at varying distances from the area of continuous stimulation. He used a vertical line of light as his continuous stimulus, and a dot of light as his test stimulus. He first of all established the intensity threshold for a dot of light at a given position. This was followed by stimulation with the vertical line for 20 sec. This in turn was followed by a re-determination of the threshold. This procedure was repeated with dots of light at different distances from the vertical line.

Two findings emerged. The first was that continuous stimulation resulted in a rise of thresholds. The second was that the rise in threshold decreased as the distance from the site of continuous stimulation increased.

IV. THE EXPERIMENT

(i) *Purpose*

To test the deduction that normal subjects would be unable, after they had been subjected to a period of continuous visual stimulation, to perceive visual apparent motion at speeds of successive stimulation which were normally fast enough to evoke such perception

¹ A re-reading of Pavlov since the completion of this experiment shows that the long-term effects of brain damage on inhibitory effects were rather to weaken than to intensify them. This poses further problems which are not relevant to the outcome of this particular paper, such as the possibility of differences between lower and higher animals.

Mrs M. Brenner was responsible for the final details of the experiment and for carrying out the testing.)

(ii) Apparatus¹

The apparatus consisted of a modified Dodge tachistoscope copied from the model given by Werner & Thuma (1942) in their article; the timing device was electronic, and had continuously variable speeds for the time interval from 0.110 to 0.86 sec. Exposure time was kept constant at 0.09 sec.

The apparent motion figures consisted of two white lines, $\frac{1}{20}$ in. wide and $\frac{5}{8}$ in. long, making an angle of 70° . The lines were made of white sellotape mounted on black velvet cards, 6×6 in., and viewed at 20 in. from the subject. The lines were so placed that the one side of the angle was horizontal, and the angle pointed to the left (see Fig. 1).

For continuous stimulation it was decided to use a figure consisting of a black circle, 2 in. diameter, mounted on a white screen 3×3 ft. It stood at eye level, 6 ft. away from the subject, so as to occupy the same visual angle as the two lines. The area of continuous stimulation was thus placed in the path of the apparent motion.

A chin rest was used during fixation, and the experiment was carried out in a semi-darkened room. To avoid communication from the after image, and to make sure we were dealing with central and not peripheral functions, the continuous stimulation and apparent motion were done monocularly, using a different eye for each—the dominant eye being used for tests of apparent motion.

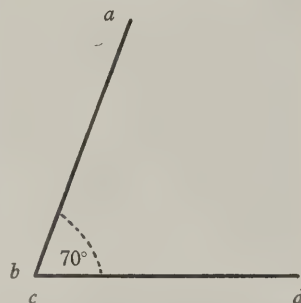


Fig. 1. Line *ab* was first exposed, and followed by line *cd*.

(iii) Sample

The subjects for this experiment were thirty-three normal intelligent adults. Of the thirty-three subjects, seven failed to report movement—so the complete experiment used twenty-six subjects who were divided into an experimental and a control group. Table 1 gives the distribution of the groups with respect to age, sex, visual acuity and sophistication in experiments of a visual perceptual nature.

Table 1. *Distribution of age, sex, vision and sophistication of two groups of subjects*

	Age		Sex		Vision			Sophistication	
	21-25	25-35	M.	F.	Normal	Corrected	Uncorrected	No experience	Some experience
Experimental	6	7	6	7	10	2	1	8	5
Control	5	8	7	6	8	5	0	9	4

(iv) Procedure

Test 1. First the threshold for apparent motion for each subject was established on five trials using the limiting method. The lower threshold was the average of the longest time intervals at which movement was reported for each of the five trials.

In preliminary trials many subjects reported difficulty in deciding whether or not they saw clear optimal movement. Subjects were therefore instructed to answer 'no' when they were quite sure that they perceived clear succession ('two separate lines one after

¹ Mr W. Withers constructed the timing device and Mr J. Westhenry the tachistoscope.

the other') and to answer 'yes' to 'any impression that the first line moved down to the place of the second'.

Test 2. A second test was carried out on each subject between 24 and 48 hr. later. In this test, the experimental group was first treated with continuous stimulation in the appropriate part of the visual field. This was done by fixating a cross in the centre of the circle described above for 3 min. The subjects kept as still as possible while doing this. Immediately afterwards, the threshold for apparent motion was established in the same way as in Test 1, with the same instructions. The time between the two procedures was not measured, but the one followed immediately upon the other in the literal sense of the word.

The subjects of the control group had their thresholds for apparent motion established without prior fixation.

It is to be noted that, although several subjects commented on the difficulty of deciding whether or not they saw movement, the thresholds appeared to be relatively consistent for the five trials of each test. This might be due to the instructions which resulted in any experience which was not definitely 'successive' being regarded as 'apparent movement'.

V. RESULTS

Table 2 gives the lower thresholds of each subject for each of the five trials in the first and the second tests. As it was not possible to produce simultaneity under the given conditions changes in the upper thresholds could not be measured and are therefore not given.

The results are consistent with expectation. The lower thresholds for the experimental group rise, i.e. after continuous stimulation the subjects fail to see apparent motion at speeds of succession which were sufficient in the first test. They now require higher speeds. The difference between the first test and the second test gives a correlated t of 3.5 which is significant at the 0.005 level. The mean lower threshold of the control group also rises, but insignificantly, giving a correlated t of 0.64. (See Table 4.)

To check the possibility of abnormal sampling influencing the result, the mean difference between the experimental and control groups at the first test was examined. The control has a mean threshold which is higher than that of the experimental group, but not significantly so.

The difference between the control group and the experimental group on the second test is even smaller and the t falls to 0.84 (Table 3). To analyse this finding further the mean threshold for each of the five trials run in each test was calculated and the results plotted in the graph in Fig. 5, p. 66.¹ Two main findings emerge.

First, in the experimental group there is a marked rise in threshold in the trial immediately after the continuous stimulation, and this threshold falls steadily except for the fourth trial which is roughly the same as the third. That this reflects a real trend is shown by the fact that the first two trials of the first test differ at the 0.005 level of significance from the first two trials of the second test. However, the last trials of each test do not differ significantly. This applies only to the experimental group. There is no such trend in the control group (Table 4).

The second finding is that the mean threshold for trial a of the second test is higher in the experimental group than in the control. The difference is significant at the 0.02 level.

¹ This was done at the suggestion of Mrs M. Brenner.

in trial *b*, which is the second trial, the difference between the two groups is not significant (Table 3).

It would seem, therefore, that the results are in every way consistent with expectation. This is confirmed by comparing the mean change from test 1 to test 2 in the experimental

Table 2. *Showing all the lower thresholds obtained in milliseconds. Blanks are shown where the subject said he saw nothing*

Control group												
Test 1. Trials							Test 2. Trials					
<i>S</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	Mean	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	Mean
1	120	400	300	300	300	284	300	120	400	120	300	248
2	400	120	200	300	300	264	300	200	200	300	300	260
3	300	150	150	150	200	190	200	200	150	300	150	220
4	110	300	120	150	120	160	200	200	200	150	150	180
5	110	200	150	200	200	172	120	200	200	200	200	184
6	200	120	150	200	150	164	200	150	150	200	150	170
7	—	150	—	150	110	136.67	200	200	150	150	150	170
8	200	150	300	200	300	230	200	200	200	300	200	220
9	200	200	150	200	200	190	120	150	120	200	200	158
10	300	200	200	300	300	260	200	200	300	200	200	220
11	200	200	120	—	—	173.33	120	120	120	120	120	120
12	500	200	400	500	400	400	400	500	300	500	500	440
13	300	200	400	200	300	280	200	300	300	200	200	240
Mean	245	199.23	220	237.5	240		212.31	210.77	214.62	233.85	216.92	

Experimental group												
Test 1. Before continuous stimulation							Test 2. After continuous stimulation					
<i>S</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	Mean	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	Mean
1	300	200	300	200	200	240	150	200	200	300	200	210
2	300	120	200	110	200	186	120	150	120	150	120	132
3	150	150	120	150	200	154	110	110	120	200	400	188
4	200	200	300	300	300	260	150	150	200	150	400	210
5	400	400	200	500	300	360	150	120	500	200	400	274
6	200	200	200	150	200	190	150	300	200	300	300	250
7	300	300	400	300	400	340	200	200	300	150	300	230
8	300	200	300	300	200	260	150	150	—	200	200	175
9	400	400	300	300	300	340	150	400	200	200	300	250
10	150	150	300	200	150	190	110	150	200	—	200	165
11	150	400	300	300	200	270	—	200	200	200	200	200
12	200	200	300	200	200	220	—	120	150	150	150	142.5
13	200	500	300	200	200	280	120	120	120	150	120	126
Mean	250	263.08	270.77	246.92	234.62		141.82	182.31	209.17	195.83	253.08	

Table 3. *Comparison of control group with experimental group*

Measures compared	Control		Experimental		Difference	<i>t</i> *	Level of significance <i>P</i>	D.F.
	Mean	S.D.	Mean	S.D.				
Means of test 1	223.38	69.19	253.08	62.74	29.70	1.10†	N.S.	24
Means of test 2	217.69	74.86	196.39	45.38	21.30	0.84	N.S.	24
Trial 'a' on test 2	212.31	77.07	141.82	24.79	70.49	2.78	<0.01	22
Trial 'b' on test 2	210.76	94.60	182.31	80.01	28.45	0.80	N.S.	24
Change from test 1 to test 2	5.69	30.63	56.69	54.97	51.00	2.81	<0.005	24

* Formula for Uncorrelated '*t*':
$$t = \frac{M_1 - M_2}{\sqrt{\left[\frac{\Sigma x_1^2 + \Sigma x_2^2}{(N_1 + N_2) - 2} \left(\frac{1}{N_1} + \frac{1}{N_2} \right) \right]}}$$

† Two-tail test; all other *t*-tests one-tail.

group with the same change in the control group. The change in the experimental group is larger and the difference significant at the 0.005 level (see Table 3).

We can conclude, therefore, that the data as they stand are consistent with the notion

that continuous stimulation has an immediate effect on threshold for the perception of apparent motion but that effects last for a short time.

VI. DISCUSSION

These findings are consistent with those reported by Detherage & Bitterman (1952) whose paper appeared while the experiment reported in this paper was in progress. They do not give details of their results, but they report an increase in the threshold of apparent motion resulting from a period of continuous stimulation, and indicate the connexion between this result and those obtained by Teuber & Bender (1949) with brain-damaged patients.

Table 4.

Measures compared	Groups	Test 1		Test 2		Difference	<i>t</i> * (correlated)	Level of significance <i>P</i>	D.F.
		Mean	S.D.	Mean	S.D.				
Test 1 and test 2	Control	223.38	69.92	217.69	74.86	5.69	0.64	N.S.	24
	Exptl.	253.08	62.74	196.39	45.38	56.69	3.57	0.005	24
Means of trials <i>a</i> + <i>b</i> of test 1 with means of trials <i>a</i> + <i>b</i> of test 2	Control	218.46	53.84	211.54	77.81	6.92	0.51	N.S.	24
	Exptl.	256.54	81.65	163.46	47.05	93.08	4.26	0.005	24
Means of trials <i>d</i> + <i>e</i> of test 1 with means of trials <i>d</i> + <i>e</i> of test 2	Control	233.69	84.97	225.38	92.78	8.31	0.55	N.S.	24
	Exptl.	240.77	73.27	224.62	58.37	16.15	0.75	N.S.	24

* Formula for Correlated '*t*':

$$t = \frac{\bar{d}}{\sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{N}}{(N_1 - 1)N_1}}}$$

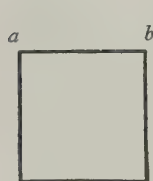


Fig. 2.

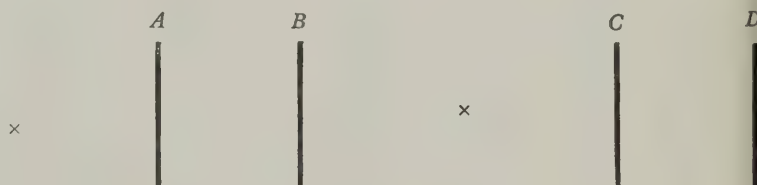


Fig. 3.

VII. ALTERNATIVE EXPLANATIONS OF THE FINDINGS

While the results reported above are consistent with the explanatory hypotheses advanced, further consideration shows that there are at least two alternative hypotheses which will account for the results as well as the irradiation theory.

(i) *Displacement effect*

One of the demonstrated effects of continuous stimulation is the displacement effect. As an example, let us assume that an area of the visual field has been under stimulation for 2 min., by a square figure (see Fig. 2). In the illustration $\times$ marks the point of fixation. At the end of the 2 min., let us assume, the square figure is removed and the two pairs of lines as shown in Fig. 3 are exposed. In this illustration $\times$ is again the fixation point. The distance between lines *A* and *B* is equal to the distance between lines *C* and *L*, the distance between line *B* and the fixation point is the same as the distance between line *C* and the fixation point. The distance between *A* and *B* is just greater than the distance between *a* and *b* of the square figure shown in Fig. 2. When Fig. 2 is exposed the distance

between *A* and *B* will appear to be greater than the distance between *C* and *D*. This is the kind of displacement effect which results from continuous stimulation. This effect was the main feature of what Kohler & Wallach (1944) called the 'figural after-effect'. It was the central point of their investigations, and was demonstrated repeatedly in different ways.

A related effect has been demonstrated in brain-damaged patients by Teuber & Bender (1949), who report that when a moving light was brought in a straight line tangentially towards an area affected by brain damage determined by perimetry, it appeared to curve away from that area. That this may be due to a displacement effect, similar to that reported by Kohler & Wallach, is shown by the work of Detherage & Bitterman (1952) with normal subjects. They found that the stroboscopic stimulus appeared to curve away from the area which had been under continuous stimulation (see Fig. 4), just as the moving light appeared to curve away from the affected area in brain-damaged patients.

When we come to examine the present experiment, it is clear that the continuous stimulation in the second test might have increased the apparent distance between the two lines. This might have been responsible for the rise in threshold of apparent motion.

(ii) *A lowered intensity of apparent stimulation*

A second effect of continuous stimulation is the one described by Bevan (1951) and referred to above, in which the luminosity was lowered by a previous period of continuous stimulation. This alone might be sufficient to raise the thresholds for apparent motion.

The formulation of these two theories shows that while we have results which are consistent with a generalized Wertheimer hypothesis, the alternative explanations still remain to be tested.

It might be thought that we have in the displacement effect an alternative to our second hypothesis: that inhibitory effects are responsible for raising the threshold for the perception of visual apparent motion. Consideration of the relevant data in the literature shows, however, that the kind of displacement effects necessary to raise the threshold have only been observed in association with inhibitory effects.

There appear to be two kinds of displacement effect. First of all there is the displacement effect which was described in section III, in which the co-existence of an inhibitory and an excitatory effect results in the displacement of the latter *away* from the former. Secondly, there is the displacement effect to be observed when, under certain conditions, two excitatory effects coincide. Under these conditions the two excitatory effects appear to be displaced *towards* each other. For example, Kohler (1940) reports that when two parts of the body are touched simultaneously, the stimulations will, under certain conditions, appear to be displaced towards each other. He reports this phenomenon for touch, vision, and sound. Thus it seems that the kind of displacement effect which will raise the threshold for apparent motion can only appear in the presence of an inhibitory effect. We do not say, at this stage, by what mechanism this displacement takes place.

In evaluating these results it must be remembered that many of the subjects had difficulty in deciding whether or not they saw clear optimal movement, and in fact seven of the thirty-three subjects tested failed to see any movement. It was decided to continue the experiment because no reason could be seen why continuous stimulation should not have the predicted effect on any form of apparent motion. On the other hand, the

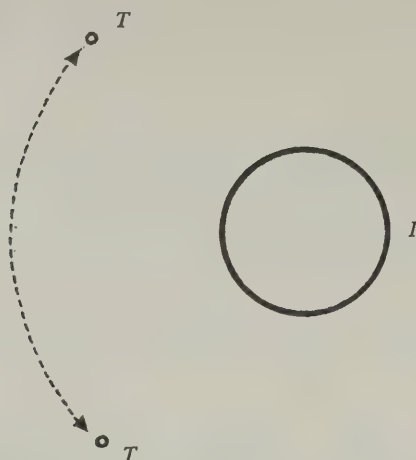


Fig. 4. Taken from Detherage & Bitterman (1952). *I*—inspection figure; *T*—test figure (alternating interocular presentation). All figures appear as light areas against a dark ground.

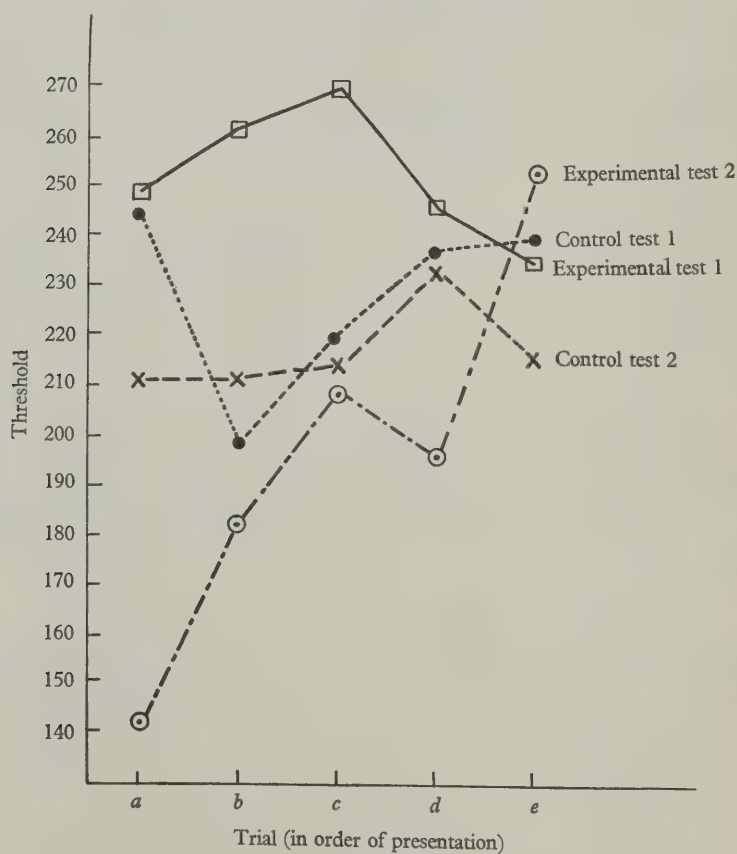


Fig. 5.

possibility still exists that where object movement is clearly seen, continuous stimulation would cease to have an effect.

VIII. SUMMARY

An explanation of the reported inability of some brain-damaged subjects to see apparent motion was advanced. This explanation was that there was an exaggeration of inhibitory effects in brain-damaged patients, and that these effects weakened or destroyed the irradiation of excitatory processes, which was postulated as the basis of the perception of apparent motion. This explanation was tested by inducing inhibitory effects by means of continuous stimulation in the appropriate area of the visual field of normal subjects. The results show a significant rise in the threshold in the experimental group, as was predicted.

Thanks are due to Mr A. Yates for the statistical analysis and criticisms, to Dr J. Briery, Senior Lecturer of the Department of Neuropathology, for discussion and advice and to Miss A. Davis for typing and preparing the manuscript. Part of the work reported in this paper was made possible by a grant from the Research Fund made available from the endowment by the Board of Governors of the Bethlem Royal Hospital and the Maudsley Hospital.

REFERENCES

- BENDER, M. & TEUBER, H. L. (1949). Disturbance in visual perception following cerebral lesions. *J. Psychol.* xxviii, 223-33.
- BEVAN, W., Jr. (1951). The influence of figural after-effects upon visual intensity thresholds. *J. Gen. Psychol.* xlv, 189-207.
- DEATHERAGE, B. H. & BITTERMAN, M. E. (1952). The effect of satiation on stroboscopic movement. *Amer. J. Psychol.* lxxv, 108-9.
- KOHLER, W. (1940). *Dynamics in Psychology*. New York: Liveright Publishing Co.
- KOHLER, W. & WALLACH, W. (1944). Figural after effects. An investigation of visual process. *Proc. Amer. Phil. Soc.* lxxxviii, 269-357.
- MARQUIS, D. G. (1935). Phylogenetic interpretation of the functions of the visual cortex. *Arch. Neurol. Psychiat.* xxxiii, 807-15.
- PAVLOV, I. (1927). *Conditioned Reflexes*. London: Oxford University Press.
- SMITH, K. R. (1948). Visual apparent motion in the absence of neural interaction. *Amer. J. Psychol.* lxi, 73-9.
- SMITH, K. U. (1940). The neural centres concerned in the mediation of apparent movement vision. *J. Exp. Psychol.* xxvi, 443-63.
- TEUBER, H. L. & BENDER, M. (1949). Alterations in pattern vision following trauma in the occipital lobes in man. *J. Gen. Psychol.* xl, 38-57.
- WERNER, H. & THUMA, A. (1942). A deficiency in the perception of apparent motion in children with brain injury. *Amer. J. Psychol.* lv, 58-67.

(Manuscript received 22 July 1953)

CORRESPONDENCE

THE EDITOR,

British Journal of Psychology.

Dear Sir,

As I have recently been supervising some research into the problems of chromatic aberration I would like to point out the relevance of this well-established phenomena as applied to Dr Zająk's paper¹ describing the results of an investigation of the influence of colour on depth perception. Some years ago I examined the effect of various colours on depth perception, but did not report the findings as they merely added to the general negative nature of the results concerning the differential effects of monochromatic light on visual acuity. However, the theory of chromatic aberration is applicable where objects in the visual field are variously illuminated with different colours. Thus, Zająk comments that 'a relation seems to exist between depth perception and wave-length of light' and 'objects viewed in indirect vision through different colour filters are seen at different distances from the eyes'.

In the second chapter of his recent text, Hartridge discusses the phenomena of chromatic stereoscopy, pointing out that if the eyes are fixated on a point the violet rays coming from that point will be most refracted and hence will appear to come from a more distant object than the red rays. Thus, if we were to align objects, in the one case comparing a red with a white object, and, in the other, a blue with a white object, then the distribution of settings in the blue would fall closer to the observer than that for the red, i.e. the blue rod, appearing further due to chromatic differential refraction, will be set nearer to the observer for equidistance, and vice versa for the red. Thus, Zająk writes that 'objects are seen as nearer with longer wave-lengths and as further away with shorter wave-lengths of light' but nowhere indicates the well-known basis of this fact.

It must be further indicated that accurate measurements will not be possible with a lateral mounting such as he has used, due to the effects of the eccentric situation of the fovea in respect of the optical axis of the eye. An experiment which used vertically mounted rods, differentially illuminated, would be a useful addition to knowledge in this field as the results could then be directly compared with those of Berry *et al.*, using the stereo-vernier apparatus suggested by Walls.

The excellent theoretical analysis of depth perception by Walls in terms of summation of subliminal monocular cues is not referred to by Zająk, but in my opinion the greatest defect in this paper is the omitting of well-established physiological findings. Experimenters in this field should also come to some agreement as to the best procedures so that results could then be compared directly and not, as at present, by analogy.

Yours faithfully,

EDGAR HOWARTH

Lecturer in Psychology

Department of Psychology

The University of Melbourne

Australia

23 June 1953

¹ Zająk, J. L. (1953). Investigations on depth perception in indirect vision. *Brit. J. Psychol.* XLIV, 132-44.

THE EDITOR,

British Journal of Psychology.

Dear Sir,

I read with interest Dr Howarth's letter to you about my paper. If I understand the meaning of his observations, his remarks refer mainly to the relation between the chromatic aberration and depth perception in *direct* and *binocular stereoscopic* vision, while the object of my investigations was the depth perception in *monocular indirect* and *binocular indirect vision with double images*. My hypothesis about differences in depth perception in monocular indirect vision resulting from differences in wave-length of light rays was advanced as early as 1913. In my recent paper I mentioned the Matthiessen formula which in my opinion is a very good illustration of the 'well-established physiological findings' as well as of other physical data mentioned by him.

I also referred to Hartridge's discussion of chromatic stereoscopy mentioned by Mr Howarth, pointing out, however, that it is connected with direct vision.

I also mentioned three different effects of refraction in the eye of pencils of light of different wave-lengths, i.e.: (1) the difference in size of the images on retinas, (2) the difference in distinctness, (3) the difference of lateral distance from the image of the fixation point; which all are results of difference of focus within the eye of an image of an object in lateral field of vision, and are related to changes of accommodation of the eye. All these effects are also undoubtedly 'physiological findings'.

I am grateful to Mr Howarth for mentioning one more theoretical analysis of depth perception by Walls, but if it refers to the stereoscopic vision, it has nothing directly to do with my investigations. Neither do the phenomena discussed by Hartridge as quoted by Mr Howarth in his letter.

I have now studied the influence of colour on depth perception in stereoscopic vision, and the preliminary results seem to confirm the fact, that, in stereoscopic vision with colours as well as without, the principal factors involved are different from those in monocular vision or in binocular vision with double images, although relations can be established between those two kinds of perception.

It must be strongly emphasized that a distinction should be made between what we *see* about depth and what we *estimate* about the distance of objects. In the case of direct monocular or binocular fixation of a single object in the field of vision there is more question of *estimating* the distance at which that object is placed than of *seeing* in depth. In my opinion it does not make sense to say that we see an object at a distance, at which it is placed in 'reality' or geometrically. But we can make estimates (rightly or wrongly) about the distance at which an object is placed geometrically. The failure of many authors to distinguish between those two notions has led to many misunderstandings about depth perception.

Mr Howarth is right to mention the difficulties involved in investigations in indirect vision and I should like to say that I am not very satisfied with my results in monocular vision, which certainly cannot be regarded as 'significant'. More significant I hope are the results with double images.

May I be permitted to mention the fact, that you, Sir, know it was necessary to cut out a good deal of theoretical discussion as well as some experimental findings of my own

and other people in order to meet the needs of the *British Journal of Psychology*. For this reason perhaps some points of my paper were not sufficiently clear.

I hope to publish some more results in the near future. These concern stereoscopic depth perception. I hope to base on them, as well as on previous investigations, one more theoretical approach to the interesting question of depth perception.

Yours faithfully

J. L. ZAJĄC

Department of Psychology

The University of Edinburgh

21 July 1953

PUBLICATIONS RECENTLY RECEIVED

Symposium on Fatigue. Edited by W. F. FLOYD and A. T. WELFORD. London: Lewis and Co. Ltd. 1953. Pp. viii + 196. 24s.

This volume includes most of the papers read at the 1952 Symposium of the Ergonomics Research Society. Since the Society publishes no journal, this public view of its activities is very welcome, and a discussion of fatigue is not only well suited to the Society's function of bringing together biologists of several disciplines, but also timely. The realization, after the first World War, that the problem of fatigue was less simple than it had at first been supposed, has been followed by a host of investigations thrust upon us by the problems of the second World War, and this is a good time to take stock of the progress made and the insights gained.

Bartlett, for example, in his opening paper, makes the good point that diminished output and increased effort, so often pointed to as criteria of fatigue, are not characteristics of performance as such, and that psychologists ought rather to look for criteria which themselves lie within the activities that are being studied. Papers by other Cambridge workers show what progress has been made in this direction.

Equally valuable to psychologists is the opportunity to take a look at current research within other disciplines. There is, for example, a good paper by Darcus from the anatomist's viewpoint, and Weston contributes a useful review of work on visual problems. Among several papers from Scandinavian authors an interesting account, by Prof. Christensen of Stockholm, of the physiological valuation of work in Lykroppa Iron Works. It is to be hoped that the Society will continue to give English readers accounts of work not readily accessible to them.

The book, which contains twenty papers in all, has been very well produced, the quality of the photographs and diagrams meriting particular mention.

N. E. LOVELESS

Divine Horsemen. The Living Gods of Haiti. By MAYA DEREN. London and New York: Thames and Hudson. 1952. Pp. x + 350. 25s.

This book is a difficult one to review adequately for the following reason. The writer states clearly that she is an artist and not a scientist and because of this it would be unfair to criticize her work from a technical standpoint. Her preface contains a claim which, in my opinion, should be given serious consideration by social psychologists and others working on deep analyses of cultural phenomena. The aim is that in investigations which involve non-verbal communication the artist is more at home than the academically trained social scientist. Miss Deren says: 'I am referring to the communication of concepts and ideas by means addressed to emotional and psychological perceptions rather than in terms of a self-conscious articulation or an address to intellectual analysis; for example the idea of power conveyed in posture rather than in the label of a name or definition.' And again: 'In effect, sensitivity to form provides the artist with a vast area of clues and data that might elude the professional anthropologist whose training emphasizes precisely that "scientific" detachment which may muffle even his normal sensitivity and responsiveness to formal nuance and subtlety, so that he becomes dependent upon the vagaries of informants' memory, intelligence and articulation.'

Miss Deren's book illustrates both the benefits and difficulties which can arise from the use of this method. On the positive side she has penetrated Voudoun to a depth which is rare in field work. On the other hand her genuine affection for the people and participation in their ceremonies has involved her to such an extent that she herself experienced possession. This involvement is perfectly understandable to anyone who has worked in the West Indies. It does, however, raise again the problem of how the investigator can participate enough to understand his people and yet retain that amount of aloofness which prevents him from being swept away by, for example, the insidious rhythm of drums.

This same involvement has made Miss Deren overemphasize the uniqueness of Voudoun as a religion. Voudoun appears to be a form of response to past and present social conditions which is prevalent with variations over the whole West Indies. The importance of her work to the scientist is that she has studied Voudoun in its most intensive form, where it is enough of an integrated institution to have a ritual which is explicable. In Jamaica anyhow the metaphysical system has become attenuated. In Haiti one is possessed by a specific loa, in Jamaica by 'the Spirit' or by that of a specific dead person. The difference

is probably accounted for by the fact that Protestantism did not provide the necessary supernatural figures of the saints.

The language in which the book is written at times becomes somewhat esoteric. In spite of this there is much in it which is of the greatest value to those workers interested in the West Indies, in religion, in folklore and to those social psychologists who are interested in people rather than abstractions.

MADELINE KERR

The Origin of Intelligence in the Child. By JEAN PIAGET, translated by MARGARET COOK
London: Routledge and Kegan Paul. Pp. xi+425. 25s.

The publication of an English edition of this important book will be generally welcomed by all psychologists concerned with the development of young children and by those concerned with the nature of intelligence and the formation of concepts. The book follows a very clear pattern, tracing the intellectual development of children from their first observed activities, an hour or so after birth until their behaviour can be said to be intelligent, somewhere about the age of eighteen months. The study is based on close observation of Prof. Piaget's own three children; there are 183 observations presented in considerable detail, some of them comparing the behaviour of all three, some tracing an activity over several days, all of them scrupulously giving the essential background of the observation and, from time to time, by reference back and forward, relating an incident to its antecedents or to subsequent behaviour to which it might be a fore-runner. Such a detailed set of systematic observations, by as skilled an observer as Piaget, is itself a valuable source of authentic material. This, however, is only the factual background for a very closely reasoned analysis of the nature of intelligence. The facts are collected systematically and presented so that there emerges a progressive pattern of development which is traced from birth through six successive stages. The starting-point is the schema of activity arising from the behaviour pattern produced by each of the primary reflex activities, sucking, grasping, etc. These schemata are extended by the assimilation into them of elements derived from the environment, the schemata being modified by accommodation to these elements. By mutual assimilation and accommodation, these schemata are grouped, through sensori-motor activity, and later regrouped to meet new situations until the final stage is reached, when some form of representation has been achieved, where invention of new means becomes possible by mental combinations instead of by active experimentation.

The final chapter of the book is devoted to a discussion of the relation of this system to the main theories of intelligence. It cannot be reconciled with the empirical-associationist or with the 'vitalistic-intellectualism' theories and, although it has some elements in common with gestalt psychology, a fundamental difference is found in the influence of acquired experience on the solution of new problems which is essential to Piaget's theory and superfluous in the evaluation of the pregnancy of 'form'. After discussing pragmatic theories, he concludes with a lucid résumé of his own theory.

The book is stimulating and some of the interpretations brilliant, both in their clarity and in the way they illuminate the path to further developments. The whole is closely argued and makes a first reading far from easy. The inferences and deductions seem to follow so directly from the observations that one wonders that they have not been so expressed before. Now that his clear statement is available in English one can look forward to a period of criticism, analysis and further evaluation of the theory by crucial testing and observation. The book merits very serious study

F. W. LAND

The Postural Development of Infant Chimpanzees. By A. H. RIESEN and E. F. KINDER.
New Haven: Yale University Press. London: Geoffrey Cumberlege, Oxford University Press. 1952. Pp. xxi+204. 32s. 6d. net.

The study of early postural development in the chimpanzee described in this book is based on a series of examinations extending over the first year of life of a number of nursery-reared chimpanzees. The experimenters used the Gesell-Thompson schedule in an attempt to determine the comparative rates of development in the human and chimpanzee infant; in addition to these tests they used one, for estimating the chimpanzee's capacity to support its weight on a rod, which was not used with the human infants. Other aspects of development were assessed at the same time and the author promises future publication of the results. Although the official schedule examination started at 4 weeks, a few earlier observations are included because in the chimpanzee several functions appear and mature in these first weeks of life which only develop later in the human infant. Moreover, like the human child, the chimpanzee is born with or acquires responses which he later loses, a few of which he shows only during the first month of life.

The comparative findings may be summed up by saying that in all the grosser locomotor and postural skills the chimpanzee is considerably in advance of the human infant, while the latter develops finer movements in advance of, and probably to a greater extent than, the chimpanzee. Structural anatomical differences seem to account for differences in preferred posture and the manner in which muscular skills are most readily developed and used. The chimpanzees were more variable in their development than Gesell and Thompson's human subjects in spite of being reared in a much more uniform environment; this probably reflects the fact that the scale was designed for the description and study of human development and may be less truly representative for the normal development of the chimpanzee. Considerable differences in behaviour were observed between the human and chimpanzee subjects; the chimpanzee infants were more nervous and were fearful of unfamiliar objects and situations, and it seems possible that the highly restricted artificial conditions under which they were reared may have affected their motional responses. In spite of the unnatural conditions of rearing, the authors report no significant differences in postural development between these nursery-reared chimpanzees and others which were reared by their mothers.

The book gives a detailed analysis of each item on the scale and a comparison for human and chimpanzee development on each item. The detail is considerable and at times a trifle confusing. Much of the material is printed in the form of tables which are not always easy to follow, the difficulty being increased by the fact that the legend is easily confused with the text. There are some excellent illustrative photographs.

One of the most interesting points suggested by the book (although the author does not discuss it) is that the precise psychological value of precocious postural development in the human infant may need to be reconsidered. The tendency among psychologists and nurse-maids to regard postural and locomotor development as correlated with intelligence appears a little unsatisfactory when one considers that the chimpanzee is more advanced in this respect than the human infant and the monkey even more precocious than the chimpanzee.

M. M. L.

The Human Senses. By F. A. GELDARD. London: Chapman and Hall, Ltd. 1953.
Pp. x+365. 40s. net.

Despite the somewhat excessive claim that an appreciation of man's senses is 'the highroad to the understanding of human nature', this is an orthodox and up-to-date survey of the physiological psychology of the sense organs. Its scope is defined by the traditional distinction of sensation from perception—a division which to the present reviewer seems sometimes to lead to rather arbitrary exclusions. For example, the basis of visual acuity is discussed, but that of the discrimination of visual depth is not, presumably because it is assigned to 'perception'.

Within these limits, the material is well presented. A notable feature is the generous treatment given to the senses other than vision and hearing. The chapters on the skin senses in particular are a considerable improvement on the thin fare usually provided in this field. It is a pity that the author did not find it possible to perform the same service for kinesthesia, which gets a bare seven pages.

N. E. LOVELESS

The Making of a Scientist. By ANNE ROE. New York: Dodd, Mead and Company. 1953.
Pp. xii+244. \$3.75.

Written for a popular audience, this book reports an inquiry into the life history (elicited in interviews) and the mental characteristics (assessed by intelligence tests, the T.A.T. and the Rorschach) of sixty-four eminent American scientists, divided into three groups, physicists, biologists and social scientists (anthropologists and psychologists). It was evidently a serious and competently conducted inquiry, notable for the co-operation that Dr Roe won from her distinguished subjects. The results are presented here in the form of summarized impressions, with a minimum of supporting data but ample illustrative quotations from the interviews. A brief chapter on thought processes suggests the possibility that social scientists rely much more on verbal and much less on visual imagery than the other scientists. But the most interesting material comes from the life histories. There is suggestive evidence that the loss of parent in childhood is commoner among these scientists than among the population at large, and Dr Roe points out the danger of generalizing too simply about the bad effects of a 'broken home'. She also found interesting differences in early experience between the social scientists and the other scientists, the former including more who had serious conflicts with their parents and more whose families thought themselves

superior to those around them; in later life they also included more who had been divorced. Irritating at times from its slightly kittenish style of writing, the book still deserves serious attention as reporting one of the very few psychological studies of eminent living people and as suggesting some interesting possibilities that would repay further inquiry.

D. W. H.

Analyzing and Predicting Juvenile Delinquency with the M.M.P.I. Edited by STARKE R. HATHAWAY and ELIO D. MONACHESI. University of Minnesota Press. 1953. Pp. viii+153. 28s.

This book is concerned with the problem of identifying the predelinquent child through the evaluation of personality characteristics: it is a collection of papers which deal with the use of the M.M.P.I. for this purpose. In the first three inquiries reported, it was demonstrated that this test differentiates delinquent and non-delinquent adolescents. Further work was then aimed at exploring the efficiency of the test for predicting delinquency, the most important study being the follow-up of over 4000 schoolchildren to whom the inventory had been administered. Certain of the scales suggest that it might be possible to anticipate some delinquency by this method; the authors conclude, however, that it is easier to predict which children will not commit offences than it is to identify those who will. Although they claim valid results with early adolescents, it is doubtful whether younger children would be capable of completing the inventory satisfactorily; as a tool in the prevention of juvenile crime, therefore, it would miss the age of peak incidence. This makes the book more of academic interest than of practical value. Secondly, in suggesting a method of prevention which earmarks children who are considered prone to delinquency, the authors do not mention the danger of the label itself, through the attitudes of those applying it, acting as a precipitant.

The book is a useful collection of data on the subject, but the findings would be more readily grasped if the scales were referred to by name, rather than by letters or numbers, a practice which involves the reader in constant referring back.

M. W.

Statistics for the Social Sciences. By T. G. CONNOLLY and W. SLUCKIN. London: The Cleaver-Hume Press. 1953. Pp. vii+154. 16s.

This book starts from the beginning and tells you just what you have to do with your raw material in order to produce the various statistics which are commonly included in presentations of experimental data in the fields of Psychology, Education and the Social Sciences. Virtually no previous statistical knowledge or experience is assumed. It embraces the usual techniques of calculating means, standard deviations, centiles, closeness of fit, normal distributions, 't' test of significance, the use of 'chi squared', sampling reliability, correlations, scatter diagrams, biserial and tetrachoric correlations, rank correlations and Spearman's foot-rule formula. Its main virtue would appear to be that it is a practical book from which actual methods can be applied directly to material available to the experimenter. The procedures to be adopted are clearly described and it should give the beginner a good start in actual practice. It would, I think, have been improved if it included some exercises and possibly an appendix containing some worked examples as patterns. It is a book for beginners which will enable them to acquire statistical experience, and it would seem to satisfy this limited purpose admirably. The chief defect is that although in a number of contexts the authors imply that there is, in fact, more to statistics than the mere calculation of a statistic, they give no hint of where a beginner should turn for more advanced guidance once the elements have been mastered. A selected bibliography would be helpful, as would even a suggestion that an experimenter might ask himself what use he intends making of his data before he begins to collect them rather than being content to tell him what he might do with them after he has gone to the trouble of collecting them. The book does fill a gap, and if a second edition is ever contemplated, the authors might bear in mind the needs of the more active readers who will want to carry their knowledge beyond the basic, uncritical, stage of routine calculation.

F. W. LAND

The Writing of Arabic Numerals. By G. G. NEILL WRIGHT. The Scottish Council for Research in Education. XXXIII. The University of London Press. 1952. Pp. xxii + 424. 10s. 6d.

As the author says, 'Arabic numerals are in everyday use... It is surprising to find how little has been published on the teaching of this indispensable skill.' The teacher who is looking for a manual of instruction, however, will not find this book wholly easy reading. In order to reach the valuable conclusions as to the best form of numerals for infant teaching and the best way to teach them, he must first digest a long though interesting section on the historical development of numerals and a second section on the author's own experiments on the writing of numerals by adults and children. The writer is obviously an enthusiast on his subject, however: although the reader may at first reel on learning that 'the numeral 1 can be written in more than eight hundred and fifty-seven million million million million ways', he soon finds himself studying with interest the numerous illustrations which show the development of numerals in a wide variety of different cultures and periods.

There are gaps in our present knowledge of how numerals have evolved. The author conceived the notion that through the study of how groups of people write numerals slowly and at speed some light might be thrown on the historical processes of development. To this end he conducted investigations with adult students, using among other methods that of serial reproduction. 'Both the historical and experimental data leave together the strong impression that every graphical sign tends to vary in a certain determinate direction.' In order to ascertain the best forms of numerals for infant teaching he carried out further experiments with schoolchildren. He gives some valuable facts on the numerals most likely to be written illegibly or interpreted ambiguously and suggests that some reconsideration should be given to the form of such numerals as, for example, 0, 7 and 2.

Readers interested in the historical development of numerals will find much to ponder over: the author's experiments will attract those concerned with the reproduction of visual material; educationists will welcome the last few chapters on the teaching of numerals in schools and the author's suggestions about method. But the volume suffers from the lack of a final chapter summarizing the work as a whole and its conclusions. G.K.

Visibility. A bibliography. Compiled by JACK WEINER and MORRIS C. LEIKIND. Edited by JACK R. GIBSON. Washington: The Library of Congress. 1952. Pp. vi + 90. 65 cents.

This bibliography of over 2000 items has sections dealing with laboratory investigations, factors such as atmospheric conditions affecting visibility in natural circumstances (including a section on underwater photometry), methods of measurement in the field, and 'engineering' applications of visibility data, including camouflage, blackout, navigation aids and signalling.

